FernUniversität in Hagen

Department of Business Administration and Economics

RESEARCH PAPERS

from the Chair of Marketing

Editor
Univ.-Prof. Dr. Rainer Olbrich

RESEARCH PAPER No. 2

C.-C. Buhr

Regularities in aggregated consumer behavior and prevention of stock-outs in retailing

Hagen 2005

Table of contents

Table of figures

Abstract

In retailing supply-side stock-outs (empty inventories) and demand-side stock-outs (empty shelves) have to be discerned, because both can occur independently from each other. They entail different kinds of damages and call for different kinds of countermeasures. The aim of this paper is to present one such countermeasure, a method that, based on a permanent analysis of incoming scanner data (e. g. market basket data), generates hints at shelves that might be empty. For this purpose selling probabilities for specific time periods are uses. The method is discussed theoretically and tried empirically with real market basket data provided by a German retailer.

1. The significance of stock-outs in retailing

Stock-outs

Stock-outs result as differences between actual demands and stored quantities of goods. Actual demand manifests itself in the quantities continously demanded by customers. How can retailers decide which quantities they want to store in future periods? These quantities are the most important foundation for specific supply management decisions, e. g. on purchase order quantities or dates.

Decisions on quantities to be stored must be based on anticipated demand: *Anticipated demand* is a *forecast* for actual demand in the future. The single most important factor to calculate such forecasts most often are indicators of actual demand in the *past*. For example, one such indicator is information on past sales. Note, that this variable must be seen as an indicator and not as the wanted information itself because past sales do not tell anything about possible unsatisfied past demand.

Stock-outs cannot be completely prevented

Retailing has to live with the problem of stock-outs. Stock-outs cannot be prevented completely since there is always a trade-off between smaller and fewer stock-outs and higher logistic costs: Since anticipated demand contains a forecast error due to uncertainty, stored quantities cannot be ensured to precisely equal actual demand in the future. Increasing stored quantities above the anticipated demand level can very well create a safety margin but even that would not all-out prevent stock-outs. It can only make stock-outs less probable, this probability decreasing with increased costs. The reason for this is that many different factors like traffic jams or walkouts in suppliers' factories not subject to logistic measures of a retailer could cause stock-outs.

On the other hand, storage, transport and other logistic costs increase with the size of such safety margins. All in all retail logistics must aim at keeping stocks as small as possible and as big as necessary at the same time: Small enough to avoid storage and other costs and big enough to prevent stock-outs.

Two kinds of damages done by stock-outs in retailing can be discerned:

Unsatisfied demand

1. Stock-outs result in unsatisfied demand: in consumers not able to buy what they came for. Hence, stock-outs are the reason for foregone sales and profit margin.

Damaged consumer attitudes

2. In the long run another result of stock-outs could very well prove to be even more important for retailers: different consumer attitudes. Depending on the importance of the product out of stock in extreme cases even a one-time stock-

out could result in some customers not coming back to the same outlet, ever. But even in cases of minor changes in shopping behavior, e. g. certain customers after a stock-out come only ones or twice per week to the outlet instead of thrice, the amount lost because of foregone *future* sales and profits can be substantial.

Difficult damage measurement

Stock-out damages cannot easily be measured. Foregone sales could at least be estimated by comparing sales amounts from periods with stock-outs with sales amounts from periods without stock-outs – given that information about stock-outs is available at all (according to a recent survey more than every second retailer does not have it, *Fisher/Raman/McClelland* 2000: 120). But the amount of damage manifest in attitude changes of unsatisfied customers cannot easily be measured. These damages could doubtless be almost prevented completely by means of safety margins in stored quantities – but at a cost. The costs due to quantities which ex-post will be known to have been dispensable, on the other hand, can easily be calculated. But each reduction in stored quantities increases the probabilities for costs due to stock-out related damages.

It appears to be certain that retailing managers will use ex-post figures of dispensable stocks to press for better sales forecasts and thus smaller surpluses.

Stock-out awareness

But this object can only be achieved if *relevant* stock-outs can be known with certainty. In retailing, this is not necessarily the case: Even when a stored quantity is known to be always above zero this does not ensure that customers may not perceive stock-outs: A shelf can be empty even if a product is still in stock! But customers cannot see and thus cannot buy the product which constitutes a relevant stock-out. This case is an important problem of retail management, because it can go undetected *and* causes a substantial damage at the same time: Unnecessary storage costs are incurred because more units could have been sold had the empty shelf been detected earlier and sales and profit margins are foregone because willing customers were prevented from spending their money. This kind of *factual* stock-out can cause supply planning being based on improper assumptions and prevent retail logistic objectives from being met.

Organization of the paper

This paper presents a method to detect this kind of stock-out, enabling retail managers to better coordinate countermeasures, and tries the method with real market basket data from food retailing. Section 2 analyzes different possibilities to use sales data in supply management. It precisely defines the notions of supply-side (empty inventories) and demand-side stock-outs (empty shelves). Section 3 sketches a new information system that permanently analyzes incoming scanner data to warn staff if irregular situations for specific products are detected. Some properties of this system and the circumstances of its practical use are theoretically discussed. Section 4 uses real scanning data from a German retailer to try out the suggested method. In section 5, some results of the theoretical discussion of section 3 are empirically verified. Section 6 sums up the findings and points out some implications for the practical use of the proposed detection system as well as possible directions for future research.

2. Supply-side stock-outs and demand-side stock-outs and the problem of detecting them

It is well know that scanner data can help to prevent stock-outs. Scanner data show past sales and thus allow forecasting future sales. Such forecasts can be used to make sure that stored quantities approximate actual demand to decrease the probability of stock-outs.

Usefulness of scanner data

Such a ‘classical’ system of sales forecasting is an *offline system* because it does not use scanner data in the moment it arises but later in time, e. g. after the outlet is closed in the evening or at weekends in cases with a weekly supply cycle.

Offline system of scanner data analysis

Such an offline system can be complemented with an *online system* that permanently looks through newly sold market baskets and automatically generates a supply order if the difference between the last known quantity stored and the number of units sold since this quantity was measured drops below a threshold. The value of the threshold might depend on the time it is used at. For example, during the morning the threshold must be higher than in the evening because the retailer wants to make sure that enough units are present to satisfy demand all day long even though newly ordered units arrive only later.

Online system of scanner data analysis

But even such a combination of an offline system for sales forecasting and an online system for automated ordering cannot solve all the problems of retail outlet supply management. Two exemplary situations show why this is: 1. For any reason no staff member realizes that a given shelf is empty. 2. A given shelf is not empty, but the units laying there cannot be seen by customers because somebody has put another product in the wrong place or has pushed the remaining units too far back. (This case of a *badly assembled* shelf will not be mentioned in the remainder of this paper but it can be substituted at all positions where *empty* shelves are mentioned: The two cases are substitutable for the present concern because both prevent customers from actually reaching the product.)

Even both systems combined are not enough

Such situations cannot be detected by the offline and online systems described above because they have other attributes than the problems these systems were designed to detect: In the first case the stock might still contain more than enough units but customers cannot reach it. In the second case the theoretical stored quantiy (last known stored quantity less number of units sold since measuring it) could as well be still above the threshold below which a new order would be issued. But even if this threshold is broken it is far from sure that the empty shelve will get *noticed* by staff.

The need to discern between supply-side and demand-side stock-outs

These considerations in mind it is easy to see that retailers must strictly discern between supply-side stock-outs (the number of stored units is zero) and demand-side stock-outs (shelves are empty). The most important reason for this is that both kinds of stock-outs do not necessarily occur simultaneously: It is absolutely feasible that a shelf is empty *although* units of the lacking product are still in stock. The matrix in figure 1 systemizes the properties of the two different types of stock-outs.

		Storage		
		empty	not empty	
Shelf	empty	**I**	**II**	demand-side stock-out
	not empty	**III**	**IV**	
		supply-side stock-out		

Fig. 1: Two types of stock-outs

The general need for better sales forecasts

The cases represented by cell IV could be problematic because stored quantities are too big which results in unnecessary costs. Like cases represented by the cells I and III they bring up the demand for better sales forecasts to enable a better planning of stock quantities. Better short-term forecasts on the level of single articles could be achieved by regression models (*Ainscough/Aronson* 1999) and especially neural networks (*Kong/Martin* 1995, *Thiesing/Vornberger* 1997, *Thiesing* et. al. 1995 and *Van Wezel/Baets* 1995). The problem of better forecasts cannot be covered in the present paper.

In contrast to that, the paper is concerned with countermeasures in cases with positive stored quantities but empty shelves which are represented by cell II. An important difference between these and cases of cell I is that for the former the charge of storage *cannot* be used as an indicator for the charge of shelves. Hence, a *quantity-oriented* surveillance of sales information, which would be enough for an online system for automated order generation, will not help to detect such cases.

Outlets without stocks

Outlets without their own stock are a special case: The shelf can either be empty or not empty. Not only discounters use such outlets but also other types of retailers who try to maximize the percentage of space used for shelves. Filling shelves from stock is replaced by daily transport from a central logistic hub. Since this business model must take demand-side stock-outs into account it necessitates an especially good sales forecast. To enable this, outlets have to record during *which periods* certain shelves are empty. This information can help to compute the number of units of an article that could have been sold without the stock-out, i. e. actual demand. As pointed out above, knowing actual demand is an important step towards good sales forecasts.

To sum up these considerations the following can be stated: Even retailers whose space configuration does not allow for the distinction between supply-side and demand-side stock-outs and thus rules out certain problems connected to this distinction have reason to be interested in learning as soon as possible whether a given shelf is empty. Since theft and other kinds of losses have to be always reckoned with, they cannot reach this goal solely by subtracting each sold unit from the last known number of units and adding new supply to it. This problem does not affect the detection method presented below.

How to deal with empy shelves?

How can retailers deal with the problem of demand-side stock-outs? How can they make sure that at least every product with a non-zero stored quantity is available on the shelf?

Strolling staff is expensive and not effective

The *classical* method to do this is having staff members permanently strolling through the outlet and checking the shelves. This method is expensive (the staff members' wages and the opportunity cost incurred by not having them do something else come into mind) and it is insecure because staff members can overlook, forget or get delayed. (Some retailers even let agents of suppliers check availability of products in outlets. This is a method that certainly is not feasible for a whole assortment.) The demand for a better method to counter demand-side stock-outs is apparent.

Basic idea of a new method

Such a method lies in a different type of online system that permanently analyzes incoming market basket data and adjusts results accordingly, just like the online system for automated order generation. The difference between the two is that this second system surveys in a *time-oriented* way. The basic idea of this method is to use the fact that an article has not been sold for a disproportionately long time as an *indicator* for the existence of a demand-side stock-out. The literature only has alluded to this possibility somewhat arcanely (cf. *Arminger* 2001 and *O. V.* 2002: 26), saying that extensive testing in retail practice had shown the method to be promising. Further references are not given. A search for publications concerned with this method also has not brought up anything.

Even useful for outlets without stocks

Time-oriented surveillance of sales has an additional dimension of use for outlets without stocks. By extrapolation the measured speed of sales it is possible to forecast *when* supply of a known size will probably be exhausted. For important basic articles of an assortment this knowledge might sometimes allow to request re-supply from the central stock in time to prevent expensive stock-outs.

3. A system for the detection of demand-side stock-outs

General description of the proposed system

But how can a time-oriented surveillance of sales be carried out through the analysis of market baskets? The basic idea of thes method to achieve this presented in this paper is the following: The incident „article *A* has been sold in one market basket“ is seen as a stochastic incident that happens with a certain frequency during a time period with a certain length. With this, the probability for the incident to happen at least once during a given length of time can be calculated. Now the proposed system would issue a warning resulting in a stuff member checking the appropriate shelf if the probability for a sale of *A* since its last actual sale exceeds a set threshold value.

This is a kind of analysis that is fundamentally different from one which would only pertain to quantities or constant time periods. The reason for that is that the analysis is not concerned with storage quantitiy or other *static* information but with a *dynamic* measurement of the exact points in time market baskets were sold at.

Results of such a time-oriented analysis of past sales data can be seen as knowlegde gained by *experience* and could thus be used to *calibrate* the detection system by specifying how big the selling probabilities for specific articles actually are.

How the system would be used

We will now illustrate the use of such a system under the assumption that it is already in place and working. It is further assumed that the system has been fed with past sales information enabling it to estimate the probability distributions it needs. Continously, every new market basket is made available to the system which analyzes it and updates its status and results accordingly. With a proper filtering mechanism implementing the threshold value, this system issues a warning when the probability for a selling that has not yet happened exceeds this value. Such a warning would be used to check the shelf it pertains to *because* the lack of selling *could* be due to the shelf being empty. A staff member who is responsable for making sure shelves are full, might receives such warnings by using a mobile telecommunications device. A warning enables him to *directly* check shelves where the probability of finding something is high. In contrast to that, a staff member with the same responsibilities but without such a detection system would not know where to check first because he would not know where the probability of finding something is higher than elsewhere.

Costs incurred and utility provided by the system

Like for all measures that shall help maximizing the result of specific organizational units also for this kind of system it has to be made sure that the costs incurred are less than the additional utility provided by it. Optimally, precisely that amount of costs should be incurred that makes the difference between additional utility and costs reach its maximum. In other words, surveillance should be extended until marginal costs equal marginal utility. It seems to be reasonable to ignore costs for the analysis of market baskets because these are essentially fixed costs that are incurred anyway, for example for storing data and for doing other kinds of analyses with it. With this assumption variable costs for the presented system equal the costs incurred for having members of staff checking the shelves for which a warning had been issued previously. These costs, in turn, are directly determined by the specified threshold: The smaller this threshold is, the more costs are incurred, because more warnings are issued and more shelves are checked. For example, when the number of warnings exceeds a certain value it could become necessary to use an additional member of staff for checking shelves.

Analysis of utility

What can be said about the *utility* of shelves being checked because of the detection system? The main aim of such checks is attained when shelves are found which are actually empty and which get filled up again faster *because* of the warning issued by the detection system. In addition to that, further utility has to be taken into account which results from staff members noticing and remedying additional problems while they are on their way to check a shelf *because* of the warning.

Aggregated utility doubtless increases with increasing checking effort, i. e. with decreasing threshold values. In the extreme case, i. e. with a threshold value of 0 %, use of the detection system would equal the classical way to detect empty shelves with the only different that *all* shelves would have to be checked *permanently*. It is assumed that during normal times most shelves for most of the time are not empty and that the damage manifesting itself in unsatisfied customers is considerably bigger than the profit of the sale would have been had it taken place (for the question how customers react to empty shelves cf. *Campo/Gijsbrechts/Nisol* 2000 and their references to the literature on stock-outs). This damage could be estimated as the profit of the average market basket multiplied by a factor *x*. Even if permanent checks were conducted for all shelves, only the relatively small number of shelves could be detected and refilled that are actually empty. Obviously, in such a situation the costs incurred for the checks would exceed the utility (prevented damage) achieved by it. Moreover, both figures would reach their respective maxima at this point.

What would happen when thresholds were increased? Ideally, the detection system would go on detecting *all* shelves which are actually empty. This means that utility would stay constant whereas costs would decrease because of fewer checks. But this stays true only until a certain value for the threshold is reached: With the threshold further increasing, the number of actually empty but undetected shelves also increases which results in lost utility.

But if it could be assumed, which seems plausible, that the warnings issued by the detection system at least sometimes reflect the real situation, increasing threshold values would result in an *increased fraction* of shelves actually empty as part of all checked shelves. Put differently, under this assumption the costs of using the system would decrease faster than the utility achieved by it.

Utility and cost curves

What can thus be said about the form of the curves for utility and costs in a coordinate system?

Utility curve. With the threshold decreasing from 100 % utility at first increases rapidly because each decrease of the threshold results in the detection of shelves that are actually empty. But the more the threshold decreases the slower utility increases because the fraction of checked shelves which are not empty increases. Seen from left to right, that is from smaller to higher thresholds, utility decreases *progressively*.

Cost curve. Costs decrease strongly from their maximum at a threshold level of 100 % because each decrease of one percentage point of the threshold probably results in the number of shelves to check decreasing by more than one percentage point, at least for higher thresholds. The nearer we come to a threshold value of 100 %, the slower control costs decrease because the absolute number of shelves to control already is rather small and therefore can only decrease by a small absolute value with each further of the threshold.

Visualization of results

Figure 2 depicts the results of the discussion about utility and costs of control by plotting curves for these variables in relation to different threshold values. In the figure *N* is the name of the utility curve and *K* is the number of the cost curve. The ordinate axis shows costs and utility in monetary units. Specific values for these are not given because they depend on the specific circumstances that the system is being used in and nothing can be said about them in general.

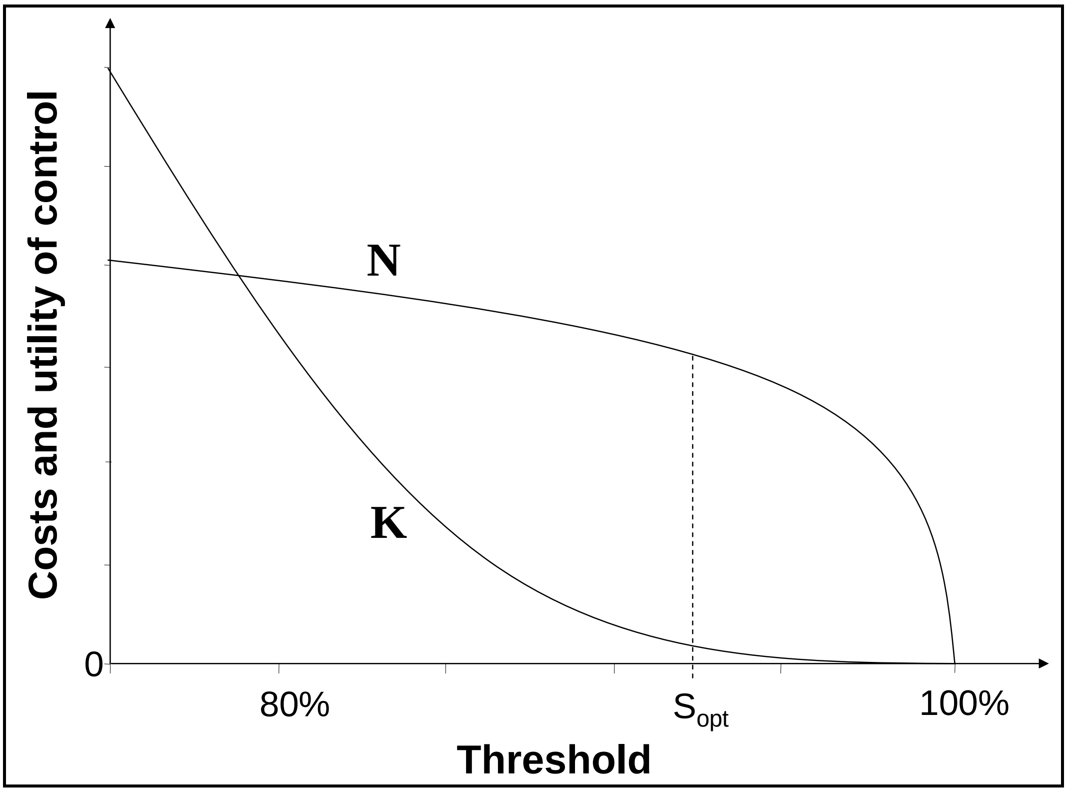

Fig. 2: Costs and utility of control and their relation to the threshold value

Area of interest between the two intersections of the two curves

If the system actually calculates what it purports (and if the detection or prevention of empty shelves makes sense at all which we shall assume here) then there must be a specific threshold value at which costs become smaller than utility before the two curves meet again at a threshold value of 100 %. In figure 2 this value can be found where the curves *K* and *N* intersect. In the interesting area between this value and a threshold of 100 % the optimal threshold value can be found where the difference between *N* and *K* reaches a maximum. This optimal threshold value is called S_{opt} in figure 1.

Threshold has to be optimized during use

It is impossible to correctly identify this optimum for a specific retailer by using analytical methods. But retailers could begin by specifying a fraction of successful checks (checks that actually find an empty shelf) which should always be reached. By way of experimenting with this fraction retailers could try and approximate the optimal threshold value.

Precise estimations for sensible threshold values cannot be given on the basis of a purely theoretical discussion as the present one has been.

4. Detection of demand-side stock-outs through analysis of market basket data

In the last section we had assumed that the detection system already existed. This assumption can now be dropped. Using some results obtained by analyzing a database of real market baskets we discuss, in an exploratory fashion, issues concerning the construction of such a system.

Real market basket data

The market basket database used for the following analysis contains a complete collection of all market baskets which were sold during the six months between November 2002 and April 2003 in four outlets of a German food retailer. All in all the database contains 1.5 million market baskets.

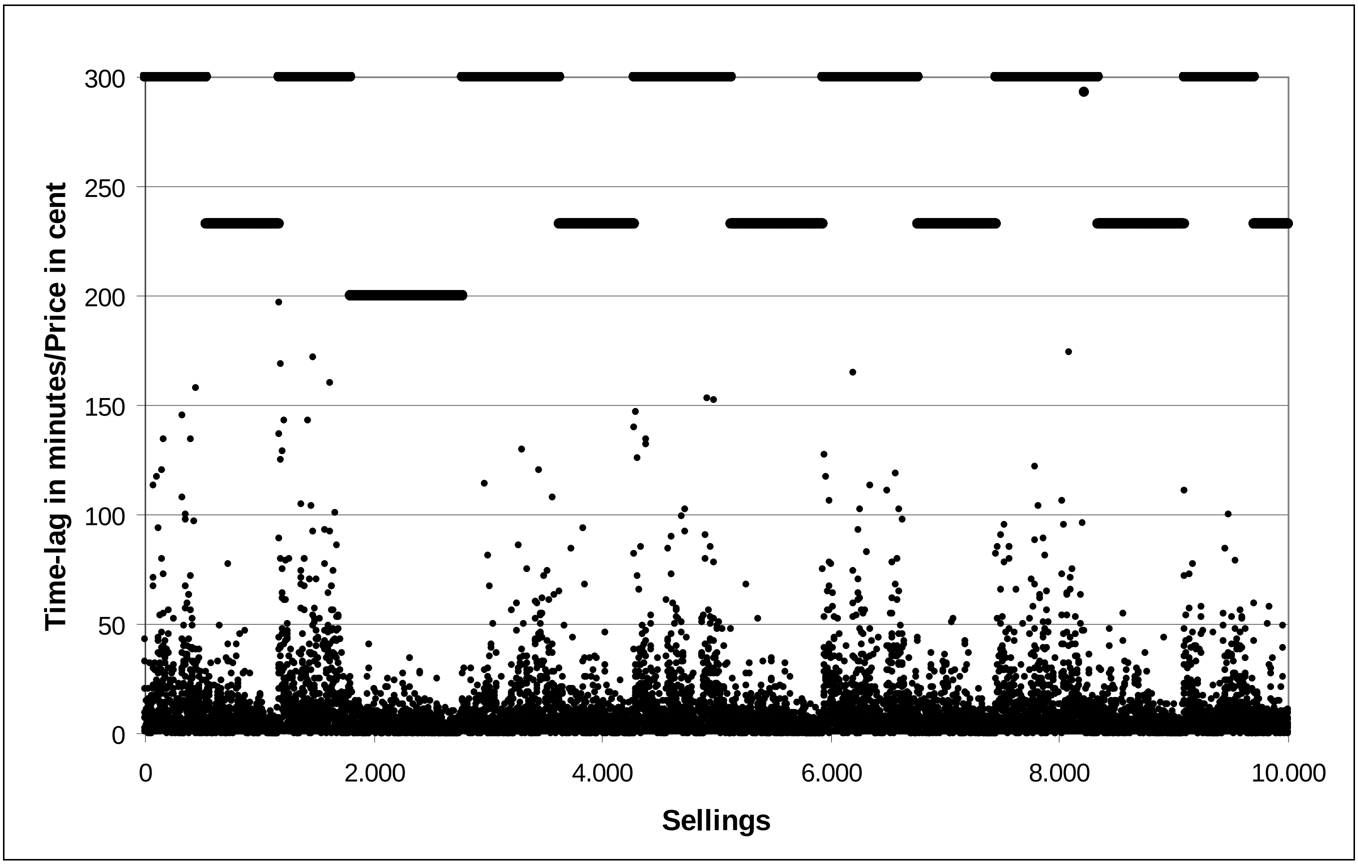

Fig. 3: Time-lags between single sellings of product *A*

Figure 3: Explanation

Figure 3 shows two time series for a specific product *A* in one outlet over the whole observation period of six months: The figure contains the prices (in cent) which were effective for each selling as well as the time lag between the given and the previous selling of *A* (in minutes). The figure abstracts from cases in which market baskets contain several units of the same product by ignoring the number of units per basket. Looking at the abscissa shows that the product *A* has been part of approximately 10.000 market baskets. Its normal price was 2.99 Euro, the most frequent promotion price was 2.33 Euro and the smallest observed price was 1.99 Euro. The time-lag for the first selling of each day is regarded as equal to the time-lag between this selling and the opening of the outlet at 8 a. m.

It is clearly visible that smaller average prices go together with smaller time-lags. Since more units can be assumed to be sold with decreasing prices this observation was to be expected. For the whole observation period the average time-lag for product *A* was 9.84 minutes and the standard deviation of all time-lags was 8.93 minutes.

Figure 4 shows the same information for a second product *B*. The price of this product has remained unchained over the whole observation period.

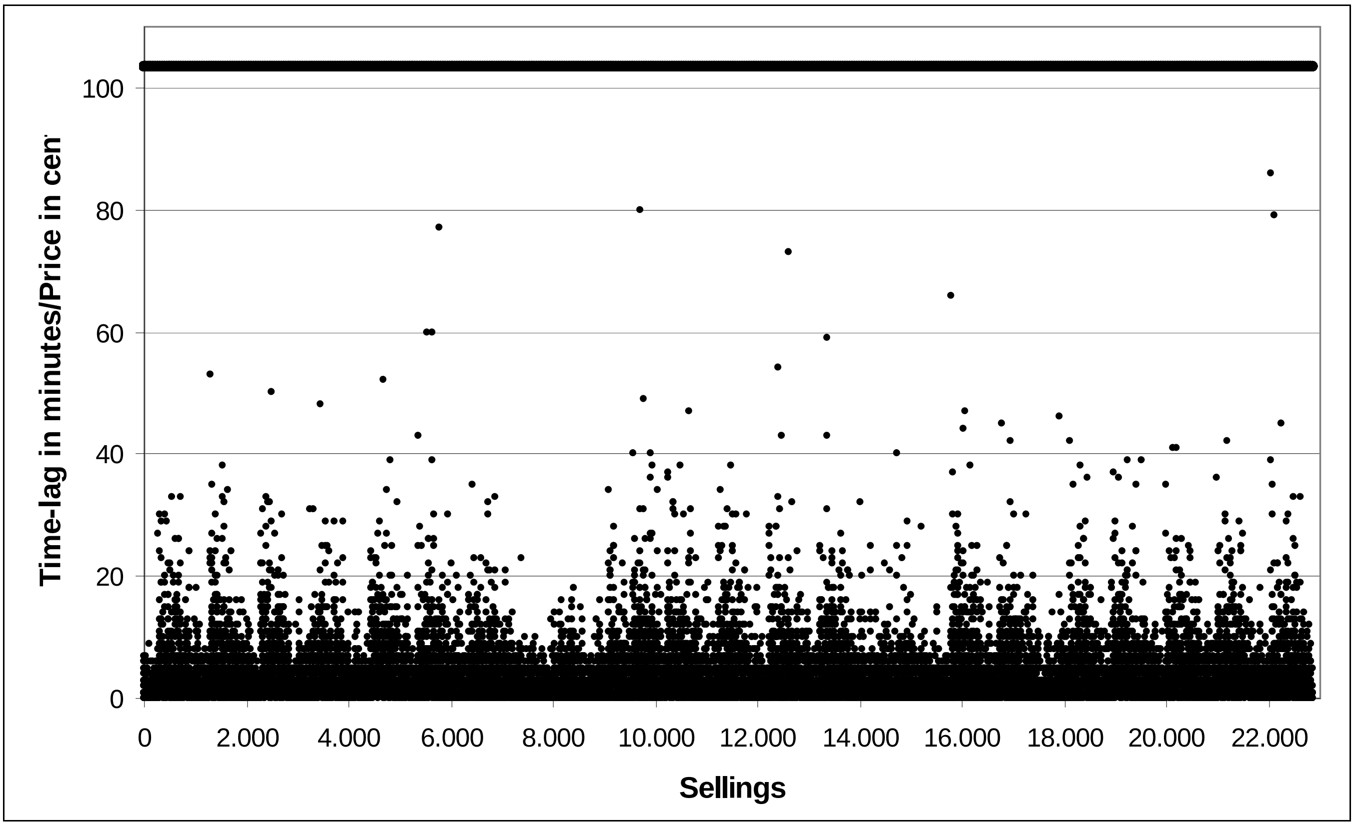

Fig. 4: Time-lags between single sellings of product *B*

In figure 3 areas with above-average time-lags are clearly associated with relatively high prices for the product *A*. Figure 4 cannot be explained in this way because prices of product *B* have been constant. Instead, a clear weekly cycle in the lengths of time-lags can be observed.

Now we transform the data from the figures 3 and 4 into another depiction: Figure 5 shows the relative frequencies of specific time-lags in black colour for product *A* and in grey colour for product *B*.

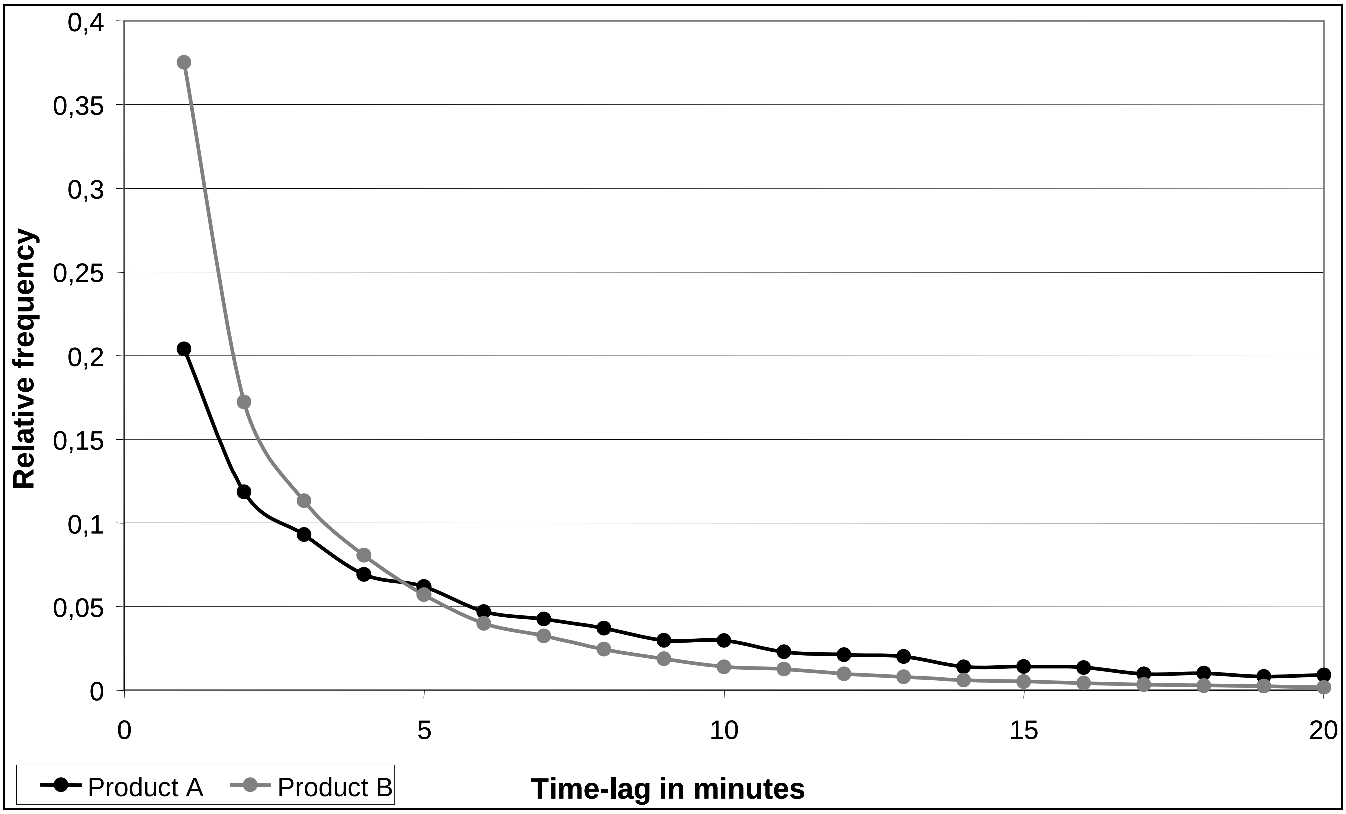

Fig. 5: Relative frequencies of specific time-lags

Figure 5: Explanation

To allow a better visibility of the important left part of the coordinate system the figure was confined to time-lags up to 20 minutes: In fact, the longest observed time-lag for product *A* was 197 minutes (for product *B* it was 86 minutes) but the frequencies of time-lags greater than 20 minutes continue to converge to zero and thus do not change the impression of the diagram. More than 87.5 % of all sellings of product *A* took place less than 20 minutes after the last a previous selling (for product *B* this fraction is 98.5 %).

Practical problem in lag measurement

A curious difficulty arose while we determined the actual time-lags from the market basket data. The 'time stamps' in the market basket database only report hours and minutes of sellings. Thus, if two baskets have the same time stamp they may have been scanned at most 59 seconds apart. But if the time stamps differ by one minute, the time-lag between the two market baskets might lie anywhere between 1 second (e. g. for 10.28:59 and 10.29:00) and 119 seconds (e. g. for 10.28:00 and 19.29:59). For time stamps differing by two minutes the actual time lag is between 61 and 179 seconds and so on.

Since it cannot be assumed that the 1-minute-difference-group contains an equal amount of sellings for each of the 119 possible actual time-lags the difference-groups cannot be adjusted and thus *overlap*. Moreover, all cases in which the two compared market baskets have the same time stamp are put into the 1-minute-difference-group because otherwise the equal-time-stamp-group would be the only one to only encompass 60 seconds. As a result we have the 1-minute-difference-group for actual time-lags between 1 and 119 seconds, the 2-minute-difference-group for actual time-lags between 61 and 179 seconds and so on.

Sales probabilities for the examples

Distributions thus derived from market basket data can be used as a basis for an online-system for the detection of demand-side stock-outs. If the distributions from figure 3 were used as estimations of the real distributions of time-lags then the system would calculate a 99 % probability for a selling of product *A* in a period of 75 minutes and for a selling of product *B* in a period of 23 minutes.

5. Conclusion, practical problems and further research

The present paper has shown why it is necessary to discern demand-side stock-outs (which are visible in the shelf) and supply-side stock-outs (which can be detected in the storage): Both kinds of stock-outs result in different damages for retailers and make different kinds of countermeasures necessary.

Different types of stock-outs

As one possible countermeasure against demand-side stock-outs we have proposed an online-system which is based on a permanent analysis of incoming market basket data. By analyzing this data in a time-oriented way this system can calculate an estimate for the probability that a specific product would have been sold since its last actual selling. If this estimation is close to the truth the system will have a quantitative cue to whether the associated shelf is actually empty. By providing a threshold for this probability the retailer can control how soon a staff member is sent to check if a shelf is empty or not.

New online detection system as countermeasure

To keep the technical demands of such a detection system as low as possible it may be worthwile to use theoretical distribution models to approximate the empirical distributions and thus reduce time and processing power needed by the system. One suitable candidate distribution may be the exponential distribution. With regard to the use of a theoretical distribution the selection of appropriate portions of market baskets for estimating the sales probabilities is particularly important. The more homogenous the circumstances were, in which these market baskets were sold, the better should be the approximation. Even without regard to the fact whether a theoretical distribution is used at all it is important to carefully select the market baskets to base the detection systems decisions on: It will be the easier to infer knowledge on specific relations in the future from data describing this relation in the past the less other factors differ between both past and future. For example, when using the system to detect stock-outs on weekend afternoon it should be calibrated with market basket data from past weekend afternoons and not with data from past weekday mornings.

Selection of data for calibration and use of theoretical distribution models

With regard to a practical use of such an online-system in retailing several problems have to be mentioned:

1. As long as it is unknown how big the damage resulting from demand-side stock-outs is, it cannot be decided whether it is promising to use such a detection system at all. To decide on the use of such a system without knowing possible

Profitablity of deployment unclear

damages for sure, estimates could be used: For each product one would need an estimate on how much profit a retailers loses each time a customer cannot find the product on the shelf. Since specific enquiries into this problem, possibly with the aid of advanced methods of consumer research, might prove to be much too expensive for the cause, subjective but conservative estimates could be used. These would probably result in an underestimation of the systems advantages and thus reduce the probability of using the system although in fact it costs more than it brings.

A heuristic way to justify the use of a detection system of the proposed kind could be the decision to use it with very high thresholds *in addition to* existent control measures. In this case costs would stay small and mostly consist in fixed costs for the development and installation of the system. The working time of staff members who control shelves in reaction to the system's warnings cannot be wholly assigned to the system because at least partly these controls happen *instead* of less productive tasks that these staff members would otherwise have spent their time with.

System should be used for critical products

2. It appears not to be advisable to use the proposed detection system for the *whole assortment* of a retailing outlet: Retailers with big assortment (several thousand different products) for most products probably never experience demand-side stock-outs. Hence, the question comes up, for which products surveillance is most urgent. Products of daily use come into mind, like milk, bread or vegetables. These articles have relatively big sales figures. Since this is an important precondition for the detection system to work properly, the proposed method seems to be able to help the most were it is needed the most.

No suggestion for the threshold value

3. It is not possible to draw a general conclusion as to how the threshold should be set at the beginning. Values below 95 % seem to be ruled out, though, because such values would entail too many warnings by the detection system. During the use of the system users get experience which will help to determine future threshold values: For example, if, for a given threshold, every warning resulted in an empty shelf being found, this would be a strong indication to decrease the threshold. Inversely, the threshold should be increased if most controlled shelves are actually not empty.

Results must be monitored to enable self-improvement of system

4. The discussion in the last paragraph shows that the proposed detection system has a tendency to improve its performance: At the beginning the system has to be based on time-lag distributions measured in the past in which systematic detection

of demand-side stock-outs was not carried out. It is plausible to assume that demand-side stock-outs have taken place in this past as well. The distributions would overestimate the average time-lag because average time-lags are longer when some customers cannot find a product in the shelf than if they find and buy it. Using such distorted distributions, the detection system at first will issue warnings later than it would using distributions measured without demand-side stock-outs.

If thresholds continue to be modified following the procedure lined out in the last paragraph, then it is to be expected that the number of actual demand-side stock-outs will decrease. Since the actual time-lags of the present will be used as an estimate for future time-lags, this estimate will improve over time. Put otherwise: The fewer the outliers in the time-lag time series (an anomalous long time-lag as a result of a demand-side stock-out would be an outlier), the easier outliers can be identified in the surveyed period.

The present paper points to further research questions in different directions.

Live tests

Most important, there is a need for trying out the system in a practical environment, i. e. to stage a test in cooperation with a retailer. Other than the data we had at our disposal the data obtained during such a test would have to include information about *actual* demand-side stock-outs. This information would help to better estimate the costs that would be associated with the introduction and use of the proposed online-system for the detection of demand-side stock-outs: It would show which frequency of demand-side stock-out is relevant in a practical environment. After the test it could be analyzed how the use of the new system has affected this frequency.

Search for good theoretical distributions and factors determining them

An important aim of further research is to search for theoretical distribution models that can approximate time-lag distributions. Particularly, for different empirical distributions different theoretical distribution models might be called for. It is an interesting question if a systematic relation between the best distribution models and characteristics of the respective empirical distributions or characteristics of the respective products exists. If such a relation was known, this knowledge could be used to identify the theoretical distribution model to be used and would avoid calibrating and comparing different theoretical distribution models for each time-lag time series.

References

AINSCOUGH, T. L./ARONSON, J. E. 1999: An empirical investigation and comparison of neural networks and regression for scanner data analysis, in: Journal of Retailing and Consumer Services, Vol. 6, No. 4, 205-217.

ARMINGER, G. 2001: Automatische Disposition und Data Warehouse, in: EHI-EuroHandelsinstitut (Hrsg.): Data Warehouse. Entscheidungsgrundlage im Handel, Enzyklopädie des Handels, Cologne, 24-27.

CAMPO, K./GIJSBRECHTS, E./NISOL, P. 2000: Towards Understanding Consumer Response to Stock-Outs, in: Journal of Retailing, Vol. 76, No. 2, 219-242.

FISHER, M. L./RAMAN, A./MCCLELLAND, A. S. 2000: Rocket Science Retailing is Almost Here. Are you Ready?, in: Harvard Business Review, July-August, 115-124.

KONG, J. H. L./MARTIN, G. M. 1995: A backpropagation neural network for sales forecasting, in: Proceedings, IEEE International Conference on neural networks, Vol. 2, 1007-1011.

O. V. 2001: Prognose und Disposition im Einzelhandel, in: IS Report, No. 7, 26-29.

THIESING, F. M./MIDDELBERG, U./VORNBERGER, O. 1995: Short term prediction of sales in supermarkets, in: Proceedings, IEEE International Conference on neural networks, Vol. 2, 1028-1031.

THIESING, F. M./VORNBERGER, O. 1997: Sales Forecasting Using Neural Networks, in: Proceedings, IEEE International Conference on neural networks, Vol. 4, 2125-2128.

VAN WEZEL, M. C./BAETS, W. R. J. 1995: Predicting market responses with a neural network: the case of fast moving consumer goods, in: Marketing Intelligence and Planning, Vol. 13, No. 7, 23-30.

The Author of the research paper

Dipl.-Kfm. Dipl.-Volksw. Carl-Christian Buhr, B.Sc., M.E.S.

born 1976,
1997-2001 Business Administration at the University of Hamburg and the Aachen Institute of Technology,
2001-2003 European studies at the Aachen Institute of Technology,
1999-2004 Computer Science and 2002-2003 Economics at the FernUniversität in Hagen,
since 2002 research assistant at the FernUniversität in Hagen (Chair of Univ.-Prof. Dr. Rainer Olbrich).

Other research papers

In German

Forschungsbericht Nr. 1:

OLBRICH, R. (1998): Handelskonzentration, FernUniversität in Hagen.

Forschungsbericht Nr. 2:

OLBRICH, R./BATTENFELD, D./GRÜNBLATT, M. (1999): Die Analyse von Scanningdaten – Methodische Grundlagen und Stand der Unternehmenspraxis, demonstriert an einem Fallbeispiel, (2., durchges. und aktual. Aufl., 2001), FernUniversität in Hagen.

Forschungsbericht Nr. 3:

OLBRICH, R./BATTENFELD, D. (2000): Komplexität aus Sicht des Marketing und der Kostenrechnung, FernUniversität in Hagen.

Forschungsbericht Nr. 4:

OLBRICH, R. (2001): Ursachen, Entwicklung und Auswirkungen der Abhängigkeitsverhältnisse zwischen Markenartikelindustrie und Handel, FernUniversität in Hagen.

Forschungsbericht Nr. 5:

GRÜNBLATT, M. (2001): Verfahren zur Analyse von Scanningdaten – Nutzenpotenziale, praktische Probleme und Entwicklungsperspektiven, FernUniversität in Hagen.

Forschungsbericht Nr. 6:

BRAUN, D. (2001): Schnittstellenmanagement zwischen Efficient Consumer Response und Handelsmarkenführung – Ergebnisse einer empirischen Untersuchung und Handlungsempfehlungen, FernUniversität in Hagen.

Forschungsbericht Nr. 7:

OLBRICH, R./WINDBERGS, TH. (2002): Marktbezogene Wirtschaftlichkeitsaspekte von Biogasanlagen nach der Verabschiedung des „Erneuerbare-Energien-Gesetz (EEG)“ – Konsequenzen für die deutsche Energie- und Entsorgungswirtschaft, FernUniversität in Hagen.

Forschungsbericht Nr. 8:

OLBRICH, R./GRÜNBLATT, M. (2003): – Projekt SCAFO – Stand der Nutzung von Scanningdaten in der Deutschen Konsumgüterwirtschaft – Ergebnisse einer empirischen Untersuchung, FernUniversität in Hagen.

Forschungsbericht Nr. 9:

OLBRICH, R./BUHR, C.-C. (2003): – Projekt SCAFO – Sortimentscontrolling im filialisierenden Handel – dargestellt am Beispiel von Frischwaren, FernUniversität in Hagen.

Forschungsbericht Nr. 10:

PEISERT, R. (2004): Die Wahl internationaler Standorte durch europäische Handelsunternehmen – Internationalisierungspfade, Strategiemuster, empirische Befunde und Handlungsempfehlungen, FernUniversität in Hagen.

Forschungsbericht Nr. 11:

OLBRICH, R./BUHR, C.-C./GREWE, G./SCHÄFER, T. (2005): Die Folgen der zunehmenden Verbreitung von Handelsmarken für den Wettbewerb und den Verbraucher, FernUniversität in Hagen.

In English

Research Paper No. 1:

OLBRICH, R./BUHR, C.-C. (2005): The impact of private labels on welfare and competition – how retailers take advantage of the prohibition of resale price maintenance in European competition law, FernUniversität in Hagen.

Research Paper No. 2:

BUHR, C.-C. (2005): Regularities in aggregated consumer behavior and prevention of stock-outs in retailing, FernUniversität in Hagen.