

ABC Analysis in an Internet Shop: A New Set of Criteria

Grzegorz Chodak, Ph.D.

Institute of Organization and Management, Wrocław University of Technology, Poland

This article presents a model of ABC analysis tailored for internet shops. The standard set of criteria is expanded to cover e-commerce specific characteristics, such as the number of product views, search engine rankings and product links via a recommendation system. The proposed new methodology is applied to real data from an internet bookstore in Poland. A comparison with the results of a standard, not internet-oriented ABC analysis shows the advantage of using the new set of criteria.

Keywords: ABC analysis, internet shop, inventory control

1. INTRODUCTION

Inventory control is one of the most important areas in internet shop management. It can decrease the costs, but also make customers more satisfied, by lowering the product delivery time. Inventory control is defined as a part of the logistics system, which concerns the material and information flow from supplier to end customer, including related activities, information systems and organisations involved [Cooper and Ellram, 1993].

From the point of view of an internet shop inventory control raises two main questions: (i) which products should be kept in-stock (if any)? (ii) how many products should be kept in-stock? The second question is the focal point of mainstream inventory theory: a shop must decide how many products to order at each point in time to meet customer demand. The first problem, on the other hand, has not been addressed in the literature too often in the context of electronic commerce (e-commerce). It involves logistics outsourcing and ABC analysis, which is the focus of this article.

E-retailing has three key features which can result in an enormous number of products in the offer:

- 2-steps shopping (the first step is on-line checkout, the second step is delivery of physical products);
- low costs of products presentation on the web site;
- easy and quick products searching.

The huge offer range needs distinguishing the important products from the set of less essential ones. ABC analysis is one of the methods which helps categorise products [Chodak et al., 2010]. The purpose is to aggregate multiple performance scores of a product with respect to different criteria into a single score for ABC inventory classification [Zhou et al., 2007]. We assume that the selection of criteria should take into account the characteristics of the e-commerce environment and the availability of data which are unobtainable in 'brick and mortar' shops. The criteria reflecting the internet information environment will be referred to later in the text as "the internet criteria". The aim of this article is to propose a set of criteria which can be used for ABC analysis in an internet shop and apply them in a weighted additive model. We provide evidence that the proposed combination of standard and internet criteria leads to different results of products classification. We also demonstrate that the proposed set of criteria can improve ABC analysis.

To meet the research objectives the contents of sections is the following. Section 2 presents shortly theoretical framework about ABC analysis. Section 3 describes general shape of the model used in internet shop ABC analysis. Section 4 presents proposed criteria applied in the model, divided into: non-internet and internet groups. In the last part of Section 4 other criteria that could be taken into consideration are shown. Section 5 presents the results of applying the proposed model in the internet book store. Finally, the last part of Section 5 discusses the obtained results.

2. ABC ANALYSIS

To conduct product classification we can use the well known ABC analysis, based on the Pareto principle. Firstly term “ABC Inventory Analysis” was used by H.F. Dickie [1951] in early 1950s [Liiv, 2007]. To have an efficient control of a huge amount of inventory items in an internet (online) shop, the traditional approach is to classify the inventory into different groups. Different inventory management policies should be then applied to distinguishable groups. The most important problem is to find products which should be always in-stock in the shop (group “A”), and differentiate the largest group of least important products (group “C”), which are often called the “long tail”. There are documented examples which show, that revenues from the long tail can be higher than from bestsellers [Anderson, 2004]. The third group, “B”, contains moderately important products which are not bestsellers, but should be analysed more carefully than the products from group “C” [Ng Wan Lung, 2007].

The interpretation of A, B, C groups in internet shop differs from traditional point of view. The ABC analysis classifies products not only for attributing the proper stock levels for different groups, but also for deciding which products should be in-stock at all [Chodak et al., 2008].

Traditional ABC analysis is based on a single measurement such as annual dollar usage. Flores and Whybark [1986] suggested that multiple criteria ABC classification can provide more complete managerial approach, allowing consideration of criteria such as lead time and criticality [Liiv, 2007]. Nowadays it is more common to take into account a greater number of criteria simultaneously. Inventory cost, part criticality, lead time, commonality, obsolescence, substitutability, number of request per year, scarcity, durability, reparability, order size requirement, stock-ability, demand distribution and stock-out penalty, are also important in inventory classification [Ramanathan, 2006]. The theoretical basis for this “multidimensional” approach is the multi-criteria inventory classification (MCIC) [Ching-Wu et al, 2008].

The proper selection of criteria should improve classification quality. It can be measured in this case as the profit level of the internet shop. When estimating classification quality we cannot analyse only one factor, like inventory cost or revenue level, because in the first case the optimal strategy will be no inventory at all and in the second all products in stock with high minimal stock values.

After building “A” and “B” group, the e-retailer should determine for these products important inventory control areas like: safety stock levels, re-order points, demand forecasting methods. These activities which follow the ABC analysis will not be the subject of our research in this article.

3. ABC ANALYSIS IN AN INTERNET SHOP – MODEL DESCRIPTION

The weighted additive model [Hastie & Tibshirani, 1990] is applied in this section for classifying inventory. There are many similar models in the literature [Ramanathan, 2006], [Ching-Wu et al, 2008], [Zhou, 2007]. The only difference in proposed model is the set of criteria which are used to carry out ABC analysis. Proposed criteria take into consideration e-commerce environment, especially data available in the internet shop.

The model is used to classify products in the best way. The objective of the classification is profit maximization by decreasing inventory costs and increasing revenues. If correct products are classified as group A, it will lead to low not-in-stock penalties, otherwise it will increase inventory costs [Chodak et al., 2008]. The key role plays the proper choice of important criteria with suitable weights.

Consider the inventory with N items and these items are to be classified based on J criteria. The measurement of the i-th item under the j-th criteria is denoted as y_{ij} . The aim is to convert the multiple measures into a single score for each of a product. Without loss of generality, we assume all criteria measures are positively related to the score of an item (if there is a negatively related criterion, transformation such as taking reciprocals should be applied beforehand).

$$S_i = \sum_{j=1}^J w_j \cdot y_{ij}$$

$$\begin{aligned}
\sum_{j=1}^J w_j &= 1 & (\text{model 1}) \\
w_j &\geq 0 \\
i &= 1, 2, \dots, N \\
j &= 1, 2, \dots, J \\
S_i &> k \Rightarrow y_i \in A \\
m \leq S_i \leq k &\Rightarrow y_i \in B \\
S_i &< m \Rightarrow y_i \in C
\end{aligned}$$

Where:

S_i is the objective function value of i -th item. A high value of this function suggests that the product should be in-stock.

w_i – weight of the i -th criteria. We assume that weights of criteria are the same for all items.

y_i – i -th item;

y_{ij} – the measurement of the i -th item under the j -th criteria;

k, m – threshold values needed to classify products.

We should transform all measures to a comparable base. We propose standard normalization.

$$y_{norm} = \frac{y_{ij} - \min_{i=1,2,\dots,N} \{y_{ij}\}}{\max_{i=1,2,\dots,N} \{y_{ij}\} - \min_{i=1,2,\dots,N} \{y_{ij}\}} \quad (1)$$

Where:

y_{norm} – normalised value of the i -th item under the j -th criteria ($y_{norm} \in [0, 1]$);

y_{ij} – the measurement of the i -th item under the j -th criteria before normalization;

$\max \{y_{ij}\}$ – maximal value of the i -th item under the j -th criteria;

$\min \{y_{ij}\}$ – minimal value of the i -th item under the j -th criteria.

Considering the weights, there are two different kinds of models. In some models the weights are endogenous. They are automatically generated when the model is optimized [Ramanathan, 2006]. In our model the weights are exogenous and the decision maker has to rank the importance of criteria. The only constraint is that the total sum of all weights should be 1. The weights obtained are common to all items in the products population.

4. DESCRIPTION OF THE CRITERIA USED IN MODEL

As we have mentioned in the Introduction, the innovative aspect of the proposed model is a new set of criteria. It consists of two groups: non-internet criteria and internet criteria. The first group takes into account the standard set of criteria, which are typically applied in traditional trade. The second group considers features of products which can be evaluated only in an internet environment.

4.1. Non-internet criteria

In literature we can find analysis, which suggest features that should be taken into account when company chooses the inventory policy in internet shop. Bailey and Rabinovich [2006] selected the following features: sales rank (merchandise popularity), release date (vintage) and price. Our list of criteria can be treated as enhancement and elaboration of mentioned features. We also point out if data needed to evaluate the criteria value is given in typical internet shop database or it requires additional software to collect the data.

Firstly we distinguish the demand driven criteria, which should be taken into consideration, when classifying products:

- time weighted sales (TWS). There are many different methods of considering the influence of time in given criteria. In our model we propose decreasing proportional weights. The sale in last period of time has weight 1. Sales in the previous periods are proportionally less important, thus the weights are decreasing by $1/(\text{number of periods})$. We used proportional weighting, but also other methods like exponential smoothing can be applied without loss of generality.

$$TWS = \sum_{k=0}^{n-1} \frac{n-k}{n} \cdot S_k \quad (2)$$

- n – number of periods;
- S_k – sale in k -th period of time.
- Data about products sales (quantities, dates) are stored in every internet shop database, thus only simple SQL query need to be built.
- the number of days after adding a product to the offer. This factor is connected to product lifecycle. We assume that an increase of selling days would decrease the number of purchases. As it was mentioned before, we assume that all criteria measures are positively related to the score of a product, so reciprocals should be applied for conversions. The importance of this criterion depends on product specification. In some sectors like electronics or computer games this criterion should be regarded as the crucial one. Date of adding a product to the offer is usually the field in table of products description, thus no additional software is required to get this data.
- deviation from average price of L internet shops which have the largest percentage share in the branch or are best positioned in the search engine which has the biggest market share. Unfortunately this data can not be explored automatically from internet shop databases, thus this criterion requires additional software development. Software single-task bots (similar to those which search engines use, but much simpler) can do this work for analytics. Data from price comparison services (PCS) were considered to be useful, especially position in the top PCS. We abandoned this idea, because usually the cheapest shops in the branch, presented on the top places in PCS are not a representative sample (e.g. the offer of the biggest Polish internet shop Merlin.pl is not presented in PCS). As all criteria measures should be positively related to the score of a product, we shift the value in such a way, that the minimal value of this criterion will be zero and use reciprocals of price deviation. The number L , which should be taken into consideration, depends on market structure. The more competitive the market is, the larger the number of e-shops that should be analysed.
- The promotion criteria could be analysed in two ways. The simplest idea is to use a Boolean value (1 – for products which are promoted in any way, 0 – for products which are not promoted). In some kinds of promotion, e.g. when an e-shop uses a sponsored link with a given budget for specific products or a group of products, the value of this criterion could be the budget for a given item. Knowing the specification of marketing activities, managers decide which kind of criteria estimation to use.

We took into consideration only one cost-oriented criterion:

- the average inventory cost for an item. As we analyse which products should be in-stock, the larger average inventory cost should decrease the objective function value, and thus we need to use reciprocal of that cost.

4.2. Internet criteria

In this section we present the criteria considering data which are available in Internet environment. It can be treated as enhancement for standard ABC analysis, useful in internet shop. We also point out the data sources needed to get product criteria value.

- The average number of product views during last D days. It is counted from a web server log which maintains a history of page requests. There are many server log analyzers available in Internet like Google Analytics, Webalizer, AWStats and others. The product views can be also taken from a database if the e-cart software enables collecting such data. More

sophisticated analyses, concerning the dynamics of product views should give better results, but need additional software to be built.

- the number of linked products, which will be presented in a recommendation system. There are many variations of a recommendation system, based on Web usage mining, product taxonomy, association rule mining, and decision tree induction [Jae Kyeong et al., 2002]. The number of linked products is not always easily available, especially when an e-shop uses a dynamic behavioral recommendation system. In some cases, when an e-shop uses simple recommendation systems like “customers who bought product, bought also...” (e.g. used in standard version of OsCommerce [www.oscommerce.com]), the simple SQL query can return the number of linked products. The greater the number of linked products, the higher the probability of demand increase for such products. Hence, this factor should be considered in ABC analysis. The more sophisticated analyse could consider not only number of linked products, but also the rank of linked products. If analysed product is linked with bestseller, the probability of demand increase is higher than if it is linked with unpopular item. Thus we propose the following function, which counts the total sell of linked products:

$$FLP = \sum_{i=1}^{NLP} LPS \quad (3)$$

Where:

FLP – value of linked products criterion – total sell of linked products,

NLP – number of liked products,

LPS – linked products sell.

- the position on the organic search results in the most popular search engine (e.g. Google) after typing in keywords describing an item, e.g. the name of the product. A position in the top 10 results (the first site list) should guarantee that the product will be seen many times and it will lead to a demand increase. The position in search engine results cannot be explored from internal databases of an e-shop and additional software for automation of acquiring this data is needed. As these criteria measure should be positively related to the score of a product, we use reciprocals of search engine position. It should be also considered other functions which assign the position on search engine results the measurement of criteria [Craswell et al., 2008]. A hyperbolic function can better represent the importance of first few positions on the search result list and the lesser importance of the lower positions.

4.3. Other important factors

The proposed list of criteria could be extended to include special factors related to the demand in a given sector. For example, in internet bookshop manager should analyse the release dates of the films based on the books offered in his bookshop. In hi-tech branches monitoring generation changes is important, e.g. the market launch of LCDs caused a rapid decrease in sales of CRT monitors [http://en.wikipedia.org/wiki/LCD_television]. In such cases it is important to lower stocks of products from the previous generation when anticipating a change.

Other factors that could be added to ABC analysis are the following:

- the number of requests in search engines, concerning a product. Such data can be collected by software created by the search engine developer - for example in Google there is an option in Google AdWords software which enables checking the popularity of keywords.
- the number of times an item is “placed” in the shopping cart without starting checkout. This number is lower than product views and suggests the potential demand for a product.
- the number of selections of an item in abandoned shopping carts. Many on-line consumers bail out of a transaction and abandon the shopping carts before the purchase transaction took place. This is very important to analyse the content of abandoned carts, because it can suggest improper inventory control, especially for those internet shops which inform about the stock availability on e-shop web site.

5. AN EMPIRICAL EXAMPLE

In this section the presented model is applied to classify products from an e-shop, which specialises in books and multimedia, mainly for children. The analysed e-shop has been operating in the Polish market for over seven years and works on OsCommerce software [www.oscommerce.com].

5.1. Presentation of results

Below the presentation of proposed model appliance is described. In Tab. 1 the values of analysed criteria are shown. The columns with proposed internet criteria were shadowed. The details concerning how analysed criteria were calculated are shown below.

- The **time weighted sales** were calculated according to equation 2. The number of selling days is taken from the database. As it is shown in the Table 1, among analysed products there are some with a long selling history and some added just 2 months earlier.
- The **deviation from average prices** in 5 stores was calculated using data available in competitive stores. As it was impossible to find 5 stores selling all 10 analysed products (only 2 e-shops had all analysed products in their offer), prices were taken from 11 e-shops which were on the top places in Google, using the names of the products as keywords.
- **“Is promoted”** criteria got the value “1”, if the product was promoted at two kinds of a promotion. Firstly products which are presented on the main (starting) site of the e-shop are considered as promoted ones. Secondly we checked in the database which products have discount offer.
- The **average number of views** was calculated as the total number of views recorded in the database divided by the number of selling days. Number of selling days was calculated using the date when products was added to the offer from table “product description”.
- The **number of linked products** was calculated as the number of items with which the analysed product is shown in the field “customers who bought product A, bought also...”. There is only one recommendation system in the analysed e-shop, thus the number of linked products was easy to calculate.
- The **positions in search engines** were analysed as the number of Google answers, when the request was the name of the item. The Google search engine was taken into consideration as it has more than a 95% share in the Polish search engines market [http://www.ranking.pl/pl/ranking/search-engines-domains.html].
- The average inventory cost was taken from the manager, who estimated the inventory cost considering the purchase price, and dimension of the product.

Table 1. Values of the eight criteria, used in the ABC analysis

Item name	Time weighted sell [item]	The number of selling days	Deviation from average prices in 5 stores [%]	Is promoted? [0 or 1]	Average number of views	Number of linked products	Position in search engines	Average inventory cost [EURO per months]
	Number of Criteria							
	1	2	3	4	5	6	7	8
Bibl. Parables	48.07	2427	3.4	0	18.1	14	1	0.50
100 Bibl. Games	127.52	1595	5.0	1	23.5	29	1	0.50
Bibl. Curiosities	101.81	630	0.6	1	22.3	41	7	1.50
Christ’s Disciple	26.50	85	4.2	1	32.0	24	2	0.50
Lord Foulgrin’s L.	0.50	89	4.7	0	3.9	6	15	0.30
Bible and Prayers	0.25	89	2.9	1	8.6	0	16	0.30
Bible CD for Child.	49.11	1903	-8.3	1	10.2	24	2	0.20
Rel. Edu.&Games	4.27	326	9.0	0	6.6	23	10	1.60
Gadu Gadu CD	5.53	510	-9.5	0	6.8	12	25	1.10
Saint’s Legends	3.00	235	5.1	0	5.1	12	6	1.60

1	Bible Parables	0.279	B	0.137	C
2	100 Bibl. Games	0.560	A	0.461	A
3	Bible Curiosities	0.453	A	0.394	B
4	Christ's Disciple	0.466	A	0.412	A
5	Lord Foulgrin's L.	0.146	C	0.224	B
6	Bible and Prayers	0.249	B	0.386	B
7	Bible CD for Child.	0.420	A	0.461	A
8	Rel. Edu.&Games	0.102	C	0.046	C
9	Gadu Gadu CD	0.215	B	0.250	B
10	Saint's Legends	0.087	C	0.063	C

5.2. Discussion

As can be seen in Table 4, the results of the proposed ABC classification differ from the results of the standard ABC analysis, which does not include "internet criteria". Note, that products #1 (*Bible Parables*) and #3 (*Bible Curiosities*) were classified in less important groups and product #5 (*Lord Foulgrin's Letters*) in a more important group. To show that the proposed classification can give better results than ABC analysis without internet criteria, we have investigated the sale of products in the three months following the moment of classification.

Position #3 (*Bible Curiosities*) is one of the strategic products for the analysed e-shop, with a great sales value. Low values of three criteria: a high inventory cost, a long history of sale and a relatively high price in the e-shop caused that this product was classified in group B, regardless of the sales. Taking into account the information that this product has many linked products, it is presented in the internet shop web page with many different products. These results in a high average number of views and high popularity (see Tab. 1). Adding these internet criteria to classical ABC analysis shifted the product's placement from group B to group A. Looking at the sales in the three months following the ABC analysis shows that *Bible Curiosities* (#3) was one of the 10 best selling products. This confirms that it should be classified in group A. As this case shows, linked products criterion and generally recommendation system is the important part of marketing environment in internet shops, thus it should be properly analysed while inventory control management.

Product #1 (*Bible Parables*) has the highest number of selling days, a high inventory cost and is not promoted. Values of the mentioned criteria caused that it was classified in the least important group C. Taking into consideration that this product is well positioned in Google, which leads to a relatively high average number of views, it should be classified in the more important group B. The sales data in the three months following the ABC analysis confirms that this product should be in-stock. This example shows that inventory control in internet shop requires taking into account data which flow from the important part of internet environment, such as search engines like Google.

Product #5 (*Lord Foulgrin's Letters*) is an example of a product which was classified to a "better" group than it "deserves". Because of a short sales history (it is a relatively new product in the e-shop) and a low inventory cost it was classified into group B, irrespective of the fact that the sales were very poor. The potential e-demand, which can be deduced from the internet criteria, indicates that this product is not popular among customers. A distant position in Google search results and a very weak average number of views shows the unimportance of this product. The sales data in the three months following the ABC analysis confirms that it should be classified in group C, rather than group B (not a single item was sold!).

The analysed small number of products cannot be treated as a representative group. However, it shows that adding "internet criteria" to standard ABC analysis improves the results in some cases. We cannot definitively decide which classification is better, but the proposed criteria have influence on demand in an internet shop and will surely provide information which can improve inventory control.

The next step in this research will be sensitivity analysis that shows which of the proposed criteria has the strongest and the weakest influence on the final results. We are also going to analyse the modifications of linked products criterion, taking into account importance of related items.

6. SUMMARY

The increasing number of internet users (especially in developing countries) results in a growing number of internet shops and increasing revenues [Chodak, 2008]. Proper inventory control is one of the most important factors influencing internet shop profitability. Taking into account internet-specific criteria, like the number of product views, search engine rankings and product links via a recommendation system, improves the results of ABC classification. Information included in those criteria helps to estimate the potential demand in an e-shop. We are aware that a more formal 'proof' would require analysing a much wider set of products over a longer period of time. Nonetheless, the presented results can be treated as a realistic assertion. Of course, further research will be devoted to 'proving' this assertion.

References

- Anderson Ch., "The Long Tail", (2004), Wired Magazine, <http://www.wired.com/wired/archive/12.10/tail.html>.
- Bailey J. P., Rabinovich E., (2006), "The adoption of inventory postponement and speculation: An empirical assessment of oligopolistic Internet retailers", *Transportation Research Part E* 42 258–271.
- Chodak G., (2008), "Evolution of polish e-commerce - main trends and perspectives.", *Enterprise, logistics and innovations in knowledge based economy / ed. by Maria Nowicka - Skowron & Ralph Lescroart. Arlon: Haute Ecole "Blaise Pascal", s. 34-42.*
- Chodak G., Ropuszynska-Surma E., (2008), "Inventory control in internet shop", *Gospodarka Materialowa i Logistyka. Issue 7, p. 17-23 (in Polish).*
- Chodak G., Latus L., Pralat E., (2010), "Inventory control, demand and forecasting and shopping methods – results of research", *Gospodarka Materialowa i Logistyka., Issue: 4, p. 2-10 (in Polish).*
- Chu, Ching-Wu; Liang, Gin-Shuh; Liao, Chien-Tseng, (2008), "Controlling inventory by combining ABC analysis and fuzzy classification", *Computers & Industrial Engineering* Volume: 55, Issue: 4, November, 2008, pp. 841-851.
- Cooper, M., Ellram, L., (1993), "Characteristics of supply chain management and the implications for purchasing and logistics strategy.", *The International Journal of Logistics Management* 4 (2), 13-24.
- Craswell N., Zoeter O., Taylor M., Ramsey B., (2008), "An experimental comparison of click position-bias models", February 2008, WSDM '08: Proceedings of the international conference on Web search and web data mining.
- Dickie H. F., (1951), "ABC Inventory Analysis Shoots for Dollars Not Pennies", *Factory Management and Maintenance*, Vol. 109, No. 7, pp. 92-94, July.
- Flores B. E., Whybark D.C., (1986), "Multiple Criteria ABC Analysis", *International Journal of Operations and Production Management*, Vol. 6, No. 3, pp. 38-46.
- Hastie T.J., Tibshirani R.J., (1990) "Generalized Additive Models", Chapman & Hall, London.
- http://en.wikipedia.org/wiki/LCD_television.
- <http://www.oscommerce.com>.
- <http://www.ranking.pl/pl/ranking/search-engines-domains.html>.
- Kim, Jae Kyeong; Cho, Yoon Ho; Kim, Woo Ju; Kim, Je Ran; Suh, Ji Hae (2002), "A personalized recommendation procedure for Internet shopping support", *Electronic Commerce Research and Applications* Volume: 1, Issue: 3-4, Autumn - Winter, pp. 301-313.
- Liiv, I., (2007), "Inventory classification enhancement with demand associations", In *Proceedings of the 2006 IEEE International Conference on Service Operations and Logistics, and Informatics*, pp.18-22.
- Ng Wan Lung, (2007), "A simple classifier for multiple criteria ABC analysis", *European Journal of Operational Research* Volume: 177, Issue: 1, February 16, pp. 344-353.
- Ramanathan R., (2006), "ABC inventory classification with multiple-criteria using weighted linear optimization", *Computers and Operations Research* 33, pp. 695–700.
- Zhou P., Fan L., (2007), "A note on multi-criteria ABC inventory classification using weighted linear optimization", *European Journal of Operational Research* 182, pp. 1488–1491.