

Mathematical Models in Banking Sector in the Context of the new Economy

Maria V. Mavri
Athens University of Economics & Business
mania@aueb.gr

Abstract

Recent advances in communication technology are changing the way that traditional banking is done. Technology has added a new dimension to the competitive pressures that are already reshaping the banking industry. The resulting changes will have a great impact firstly on the reconfiguration of a bank branches network according to the dictates of the market, secondly on the design of new products and on the development and the use of alternative distribution channels and finally on the customers' switching behavior..

In this study that presents an overview of my PhD thesis, different issues relating to the reorganization of banking industry are examined. Specifically the performance of a bank branch network, the online banking as an alternative distribution channel and the duration of the relationship among customers (individuals or enterprises) are the three interrelated parties of this thesis. Optimization techniques, generalized linear models and proportional hazard models are used in order to determine products and services which are offered to bank clients through traditional and Internet channels, forecasting customers' attitudes (adoption or rejection) to the new products and the time horizon of their cooperation with their bank.

All three sections are exploring issues that put in the fundamentals of Customer Relationship Management (CRM) initiatives. Bank's management team could use the findings of this study, in order to determine specific attributes in designing financial services and products, which would add in customers' satisfaction. The proposed approach could have significant implications for enlarging the duration of the relationship among customer and bank and for maximizing bank's performance.

1. Introduction

In today's retail banking environment, where a more sophisticated consumer with less bank loyalty is becoming the norm, customer service quality is an indispensable competitive strategy. Furthermore, the stiff competition and the compression of the interest rates, forces banks to set up and put into

effect all necessary decision support systems that will enable them to dynamically plan new locations, evaluate their performance, forecast customers' attitude to new offered products and services, estimate clients' switching behaviour, and finally provide marketing support to their geographically separate units.

Bank branch construction has slowed sensitive not only to changing technology and changing customer needs, but also to a more volatile economy. As a consequence of deregulation and information technology banks offer new services and products (Burton, 1990) A branch of the future must be framed within a flexible organization and can be changed within minimal human and resource costs in order to keep pace with ever-changing customer needs (Mountinho *et al*, 1997). Since branches occupy the key position in the bank's organization, its performance is in the core of strategic directions of each bank. The relation between branch and community's performance must also be explored in order to identify the optimal scale size for a network of existing and new branches

The explosion of Internet usage and the huge funding initiatives in electronic banking have drawn the attention of researchers towards Internet banking. The new banking business relies on electronic delivery of its products and services, supported by information systems and telecommunications (Hempel, G., & Simonson, D., 1998). Although millions of dollars have been spent on building Internet banking systems, reports have shown that potential users may not use the systems in spite of their availability. This points out the need for research to identify the factors that determine acceptance of Internet banking by the users. (Luarn, P.,& Lin, H., 2005))

Moreover, customer satisfaction is recognized as being highly associated with customer value and with product price; whereas service quality is not generally considered to be dependent upon price. The more satisfied the customers, the more tolerant to price increases they are likely to be, thus resulting in greater profits (Anderson, E., Fornell, C., & Lehmann, D., 1994; Garvin, 1988).

According to Berry (1995) the transition from transaction to relationship-based marketing is inextricably linked with the increased role of quality and satisfaction given in services.

The remainder of the paper is organized as follows: Section 2 provides a brief overview of the existing literature on performance measurement; the mathematical formulation of the branch network and modular service optimization model and outlines a new iterative procedure for solving the model. Section 3 summarizes existing research on e-banking, describes the used generalized linear model, provides the hypotheses under examination and presents the results of the analysis. Section 4 describes the problem and the hypothesis that would be considered under examination by using survival functions, life tables and the proportional hazard model. Last section summarizes the conclusions of this work and suggests further research.

2. Performance Measurement and techniques of optimization

2.1 A brief literature review

The literature addressing branch performance measuring is large, but mainly focuses on management accounting approaches (e.g. Hansen, 1990; Smith & Schweikart, 1992; Thygerson, 1991), utilizing econometric models (Clawson, 1974; Doyle *et al.*, 1979), multiple regression analysis (Heald, 1972; Boufounou, 1995), Data Envelopment Analysis (Camanho *et al.*, 1999; Vassiloglou & Giokas, 1990; Avkiran, 1999) and artificial neural network-based approaches (Athanasopoulos & Curram, 1996). On the subject of branch profitability, little research has been conducted to-date (Avkiran, 1997). Below, we provide a brief overview of the existing approaches.

Traditional measures of bank profitability, ROA and ROE, are beset with problems of allocating assets, equity and net income when applied at branch level (Smith & Schweikart, 1992).

During the 1990s Data Envelopment Analysis (DEA) has been used extensively to evaluate banking institutions. The relation between efficiency and profits was first addressed by Oral (1990, 1992) through two DEA models for analyzing both efficiency and profitability. The profitability model consisted of a desegregation of expenses and income, which were considered as inputs and outputs, respectively. However, the implication of using such models instead of a usual profitability measure was not discussed. Drake and Howcroft (1994) correlated the DEA technical

efficiency score with cost-income ratios. Their results indicated that more efficient branches had lower cost-income ratios. Another study by Schaffnit *et al.* (1997) concluded that branch efficiency has a very clear positive effect on profit. Recent publications determine the performance goals using data analysis in order to define the factors or variables that influence performance measures of a bank branch (e.g., Avkiran, 1999; Hartman, 1999). Frei & Harker (1999) present a methodology that determines the role of design in calculating the efficiency of service delivery processes. The efficiency of these processes is determined by using a variation of frontier estimation (data envelopment analysis [DEA]-like) techniques. The proposed methodology replies to the question of how much inefficiency is due to process-design choice

The evaluation of the performance of a bank branch has been widely addressed, considering fixed branch networks and offering models for evaluating their performance. Most of the approaches set performance goals without considering overall performance measures, and examine a very large number of variables, some of which are not statistically significant; this limits the applicability of the models to small-scale examples and prohibits the application to practical cases where the number of branches can be in the order of hundreds. All the models developed to-date address the traditional banking sector where most services and products are offered at all locations, and as a result they do not take into account the various modular services that one branch could offer, or the overall branch network.

2.2 Problem Description

To concentrate on the problem we address in this study, consider a geographical area where bank branches are dispersed. The wide diversity of the types of customers served and the unequal competition conditions that the branches face in the different sub-areas enforce top executives to set standards or norms in order to measure the performance of each branch and satisfy in a more efficient way customer needs. The demographical, socio-economical and market data of the area under consideration are known, while management goals and strategic plans are also set, as is the number of employees necessary to accomplish customers' requirements. In order to meet management goals, evaluation and rationalization of the current branch network must be conducted, especially if the bank's strategy has focused on business growth through the acquisition of other financial institutions and the reengineering of business processes.

The goal is to reach the optimum number of branches, given a certain level of operational and fixed costs, and the optimum mix of services that each should provide. This combined objective can be linked to the weighted sum of the performance of each individual branch within the area under examination. Metrics may include: deposit balances per branch, number of new accounts per month, average of retail lending balances, number of new lending accounts, balances of financial investments, etc. The value of the metrics is a function of various internal and external (to the bank) factors, which are the problem variables. Internal factors may include: a) number of employees, e.g. tellers, corporate representatives, receptionists, administrators, etc.; b) number and level of the management staff; and c) operational variables, e.g. number of ATM's or transactions. External factors may include: area population and growth rate, average family income, number of small/corporate business establishments etc.

2.3.1 Mathematical Model for Branch Optimization

Performance Variables:

Let n be the number of branches and i their index. The efficiency of a branch involves the identification of performance variables, which reflect the corporate objectives and strategies. For branch i , we model these variables by vector $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T$ of order m (number of variables per branch), where (T) denotes the transpose operand. Each element of vector y_i represents one performance variable for branch i . The overall performance of branch i (denoted by P_i) is defined as follows:

$$P_i = \underline{w}^T \times \underline{y}^i = [w_1^i, w_2^i, \dots, w_{m-1}^i, w_m^i] \times \begin{bmatrix} y_1^i \\ y_2^i \\ \vdots \\ y_m^i \end{bmatrix} \quad (1)$$

In (1) w is a vector of dimension m representing the weights w_{ki} of the k -th performance variable, y_{ik} , of branch i ; k is the index of performance variables ($k=1, \dots, m$).

Internal and External Factors

As discussed above, branch efficiency is related to external (non-controllable) factors and internal (controllable) factors. We model these factors, for each branch i , using the following 1-dimension vector $\underline{x}^i = [x_1^i, x_2^i, \dots, x_l^i]^T$. Each element x_j^i of \underline{x}^i represents one external or internal factor for branch i . Note that j is the index of factors.

Relationships

Each performance variable (i.e., element y_{ik} of vector y_i) is related to some internal and external factors x_{ij} , relevant to branch i ; given the data concerning the initial (current) branch configuration, we can determine these relationships through regression analysis, and obtain the regression coefficients a_{kj}^i . Equation (2) expresses the relationships:

$$\begin{bmatrix} y_1^i \\ y_2^i \\ \vdots \\ y_m^i \end{bmatrix} = \begin{bmatrix} a_{11}^i & a_{12}^i & \dots & a_{1l}^i \\ a_{21}^i & a_{22}^i & \dots & a_{2l}^i \\ \dots & \dots & \dots & \dots \\ a_{m1}^i & a_{m2}^i & \dots & a_{ml}^i \end{bmatrix} \times \begin{bmatrix} x_1^i \\ x_2^i \\ \vdots \\ x_l^i \end{bmatrix} + \begin{bmatrix} c_1^i \\ c_2^i \\ \vdots \\ c_m^i \end{bmatrix} \Leftrightarrow \underline{y}^i = A \times \underline{x}^i + \underline{C}^i \quad (2)$$

In (2), A is an $m \times l$ matrix incorporating all regression coefficients that relate performance variables and internal/external factors, while C_i is the vector of constant terms resulting from the regression analysis.

The different internal and external factors have implicit relationships that can also be determined via regression analysis on the initial data. Eq. (3) provides this relationship:

$$\begin{bmatrix} x_1^i \\ x_2^i \\ \vdots \\ x_l^i \end{bmatrix} = \begin{bmatrix} 0 & b_{12}^i & \dots & b_{1l}^i \\ b_{21}^i & 0 & \dots & b_{2l}^i \\ \dots & \dots & 0 & \dots \\ b_{l1}^i & b_{l2}^i & \dots & 0 \end{bmatrix} \times \begin{bmatrix} x_1^i \\ x_2^i \\ \vdots \\ x_l^i \end{bmatrix} + \begin{bmatrix} d_1^i \\ d_2^i \\ \vdots \\ d_m^i \end{bmatrix} \Leftrightarrow (I - B) \times \underline{x}^i = \underline{D}^i \quad (3)$$

In (3), I is the unit $l \times l$ square matrix, and B is an $l \times l$ square matrix with zero diagonal terms (no regression coefficients between x_{ij} and itself). Each b -coefficient is determined by regressing the x_{ij} factor on the remaining ones. Some of the b 's may be zero, denoting a non-existing relationship between two factors. Finally, D_i is a vector of constants resulting from the regression analysis.

A necessary condition for better performance measurement is the determination of the relationships between the values of the same group of factors, among different branches. Inequality (4) depicts such a relationship for factor j :

$$[z_j^1, z_j^2, \dots, z_j^n] \times \begin{bmatrix} x_j^1 \\ x_j^2 \\ \vdots \\ x_j^n \end{bmatrix} \leq 0 \Leftrightarrow \underline{z}_j^T \times (x_j^i) \leq 0 \quad (4)$$

In (4), z_j is an n -dimensional vector, each element of which is 1, -1, or 0, and $(x_j^i) = [x_j^1, x_j^2, \dots, x_j^n]^T$. Inequalities (4) come by inferring relationships through the examination of appropriate data sets, and the explanation is as follows: We observe data concerning a specific set of factors for different branches. The data imply some relationships between these factors. The relationships are

captured in the form of constraints to the model. If such constraints are not established, the results of a solution approach may be questionable.

Inequality (4) can be expressed for each factor j ; e.g., if x_{i5} is the number of tellers for branch i , and we know that branch 2, by design, must have a smaller number of tellers than branch 3, we can write $x_{25}-x_{35}\leq 0$, implying that $z_{25}=1$ and $z_{35}=-1$.

Bounds

The elements, x_{ij} , of vector x_i may be bounded by known parameters L_{ij} and U_{ij} , which are imposed by management or external conditions. Such relationships can be expressed as follows:

$$\underline{L}^i \leq \underline{x}^i \leq \underline{U}^i \quad (5)$$

In (5), $L^i=(L_{ij})$ and $U^i=(U_{ij})$ are l -dimension vectors incorporating bounding parameters. Inequality (5) can be expressed for all the factors, since both controllable and non-controllable factors may be bounded for branch operation.

2.3.1.1 Linear Programming Formulation

Given all the above definitions and relationships, we can proceed to the formulation of a linear program to maximize the overall performance for the geographical area under examination. The model is as follows:

$$\begin{aligned} & \text{Maximize} && \sum_{i=1}^n \underline{w}^T \times \underline{y}^i \\ & && (6) \\ & \text{Subject to:} && \\ & \underline{y}^i = A \times \underline{x}^i + \underline{C}^i && \forall i=1, 2, \dots, n \quad (7) \\ & \underline{x}^i = B \times \underline{x}^i + \underline{D}^i && \forall i=1, 2, \dots, n \quad (8) \\ & \underline{z}_j^T \times (x_j^i) \leq 0 && \forall j=1, 2, \dots, l \quad (9) \\ & \underline{L}^i \leq \underline{x}^i \leq \underline{U}^i && \forall i=1, 2, \dots, n \quad (10) \end{aligned}$$

$$\underline{y}^i \in R_+^m, \quad \forall i=1, 2, \dots, n \quad (11)$$

$$\underline{x}^i \in R_+^l, \quad \forall i=1, 2, \dots, n \quad (12)$$

The model of (6)-(12) is a typical linear program, which is appropriate for measuring both the performance of a bank's network and of each individual branch separately. To recap, the objective function (6) sums up all the individual branch performances. Equation (7) relates performance variables and factors via regression analysis coefficients. Equation (8) correlates branch factors through regression analysis coefficients.

Inequality (9) correlates different branches for the same factor. Inequality (10) bounds controllable or non-controllable factors. Finally, constraints (11) and (12) force the problem variables to assume non-negative real values.

3.2.2 Solution algorithm

Up to this point, we have provided a model for bank's branch performance by integrating internal and external factors with performance variables. Bank's management team determines which performance variables and factors will be under examination according to the strategy that had already been decided to be applied. Although strategic plans cannot be quantified, what we really can measure, is some parameters crucial for the development and the application of the plan. Using mathematical techniques we can model and formulated factors (find optimum value, or a set of possible values for each factor) and to make a prediction – estimation for the applicability and usefulness of the proposed plan.

In this study we developed an algorithm called PERFORMANCE to solve the optimization problem of (6)-(12). To eliminate a branch we measure the performance of all branches and we select the one with the lowest performance score. By resolving the problem, we calculate the new performance of the reconfigured network. If the difference O_t between the overall performance and the operating costs (i.e. staff's salaries, rent for buildings etc.) for all branches increases, the procedure is repeated until $O_{t+1} < O_t$. In this case the network has reached the point of operational efficiency. Note that t is an index of iterations, which is identical to the number of branches eliminated through the iterative procedure, and $O_t = \sum_{i=1, \dots, n} [P_i - (\text{Operating Cost})_i]$. The Steps of the algorithm are provided below:

Algorithm PERFORMANCE

- Step 0: Input: number (n) of branches, number and values of performance variables y_i (branch data), and current values of the external and internal variables x^i .
- Step 1: Regress performance variables on external (non-controllable) and internal (controllable) variables, to determine matrix A and vector C_i in Equation (2).
- Step 2: Regress external and internal variables on themselves to determine matrix B and vector D_i in Equation (3).
- Step 3: Observe the value of the same internal or external variable among different

- branches and form inequality (4) with appropriate z_j .
- Step 4: Impose the bounds L_i and U_i of inequality (5) for each variable.
- Step 5: Calculate the performance of each branch separately P_i and of the whole bank branch network from (1), before optimization.
- Step 6: Solve the linear programming problem of (6)-(12) to determine the overall optimum performance for the bank branch network and of each branch.
- Step 7: Branch Elimination Procedure
- Step 7.1: Set $t=1$.
- Step 7.2: Calculate the difference O_t
- Step 7.3: Find branch i^* with lowest performance P_i .
- Step 7.4: Eliminate branch i^* and reformulate problem (6)-(12).
- Step 7.5: Resolve the problem without the branch i^* .
- Step 7.6: Calculate the difference O_{t+1} for the $n-t$ branches.
- Step 7.7: IF $O_{t+1} > O_t$ THEN a new elimination is suggested; set $t=t+1$ and RETURN to Step 7.3; OTHERWISE, keep all branches in the bank branch network since it operates effectively, and GOTO Step 7.8.
- Step 7.8: Terminate – Efficient branch network reached.

The algorithm terminates by providing the target number of branches in the catchment area, the mix of services that define the optimum performance for the bank branch network, and the operational parameters of each branch (through the values of the controllable factors). Note that we include operational costs of branches in the iterative procedure in order to ensure that branch eliminations are preferred, which is the case in the majority of branch network reconfigurations.

The proposed algorithm could be re-applied in any network, which is under evaluation and reconfiguration. Its main advantage is that the performance and factorial variables, which are under examination, are defined each time by the management team. All necessary coefficients that the algorithm uses are calculated through the update values of the above variables. So the algorithm becomes a tool that is always up-date according to the dictates of the market.

3. Electronic delivery channels

3.1 A brief literature review

Recent advances in communication technology, including the development of more powerful computers, are paving the way for new banking products and services, changing the way that traditional banking is done. The resulting changes will have a great impact on the development and use of alternative distribution channels. The most recent delivery channel is online banking. Electronic or online banking is the newest delivery channel to be offered by retail banks in many developed countries and there is a wide agreement that this channel will have a significant impact on the market. Banks know that the Internet opens up new horizons for them and moves them from local to global frontiers.

Customer adoption is a recognised dilemma for the strategic plans of financial institutions. Several studies have investigated why individuals choose a specific bank. Important consumer selection factors include: convenience, service facilities, reputation, and interest rates (Kennington *et al.*, 1996; Zineldin, 1996). According to Delvin (1995) customers have less time to spend on activities such as visiting a bank and therefore want a higher degree of convenience and accessibility.

Liao and Cheung, (2002), employed survey data and regression analysis to measure consumer attitudes toward Internet based e-retail banking as a financial innovation. They found that individual expectations regarding accuracy, security, network speed, user-friendliness, and user involvement and convenience, were the most important quality attributes in the perceived usefulness of Internet-based e-retail banking.

Many studies in the literature extended the Technology Acceptance Model which is a theoretical framework, which identifies the perceived ease of use and perceived usefulness as the key reasons for using Internet Banking. Luarn, P., & Lin, H., (2005) extended TAM and they added "perceived credibility", "perceived self-efficacy" and "perceived financial cost" to the model. Bomil Suh & Ingoo Han, (2002) in their study introduced trust as another belief in TAM that has an impact on the acceptance of Internet banking. Wang *et al.*, (2003) also used the TAM and they introduce "perceived credibility" as a new factor that reflects the user's security and privacy concerns in the acceptance of Internet banking. The study also examines the effect of computer self-efficacy on the intention to use Internet banking. It also

demonstrates the significant effect of computer self-efficacy on behavioural intention through perceived ease of use, perceived usefulness, and perceived credibility.

Koufaris, *et al* (2002) examined the impact of consumer experience and attitudes on intention to return and unplanned purchases on line. Kambil *et al*, (2000), show that senior management's support and technical issues such as information security are of the most significant impacts to firms that take their business online.

Mols, (1999), determines that bank customers are divided into Internet banking segment and a branch banking segment. The Internet influences the future distribution channel structure in two ways: (a) it is in itself a new distribution channel for financial services and (b) it influences consumers in a way that they invest time and resources in becoming PC-literate and in familiarizing themselves with the Internet. Users of PC banking are more satisfied, are less price sensitive, have higher intentions to repurchase and provide more positive word-of-mouth than non-users.

Hitt and Frei, (2002), examine whether and how characteristics or behaviors might differ between customers who use electronic delivery systems and those who use traditional channels. By using logistic regression they conclude that demographic characteristics and changes in customer behavior following adoption of PC banking account only a small fraction of overall differences. Karjaluoto, *et al*. (2002) explored the effect of different factors affecting attitude formation towards Internet Banking in Finland. By using factor analysis they determined that prior experience of computers and technology as well as demographic factors impact heavily consumers' online behaviour. Jun and Cai, (2001) used the Critical Incident Technique (CIT) to uncover the key dimensions of Internet banking customers and to identify critical satisfying and dissatisfying factors which were reliability, responsiveness, access and accuracy.

3.2 Problem Description

The purpose of is to present a methodology for identifying factors that affect someone's decision using or not using online services. In order to meet this goal we use a generalised linear model, specifically a logistic regression model, and we are trying to estimate the factors that improve someone's final decision and the contribution of each one. The scope of this study is to describe bank-customers' behaviour in Greek online market, which is fragmented and Internet adoption within the population is quite low, up to 22,4% due to a

survey that is completed by National Statistical Organization in 2005, a fact that offers a basis for initial estimations.

The study therefore, focused on identifying the significance of the following factors:

H₁: Speed of transaction of electronic delivery channels provides a competitive advantage for them.

H₂: Difficulties in the use of the new technology prevent some customers from using it.

H₃: Many people believe that today's "traditional" banking system operates well and thus, the online presence of the banks is characterized as not necessary.

H₄: Internet banking costs include those associated with Internet activities as well as bank costs and charges. Cost influences consumers' attitudes toward electronic services.

H₅ (a): According to literature, young people, from 20 up to 40 years old, are familiar to Internet use

H₅ (b): People who use electronic banking services have a higher education level than others. Maybe education is another factor that is characterized as important for using or not using online services.

H₆: Many people complain about lack of information concerning the new electronic channels that financial institutions use today.

The above hypotheses are tested due a field survey questionnaire that took place in Athens and Thessalonica during March of 2002. Findings of the questionnaires show that:

- The competitive advantage of e-banks or electronic banking services is the speed of the transactions. More than 90% of the respondents consider it, to be a valuable attribute in the electronic financial industry. Avoiding queues or delays are the primary reason for choosing online transactions.
- It is useful for the banks to advertise their online presence, the services and the facilities that they offer to the clients; 89% of the respondents ascribe deficiency of information concerning the electronic banking services as the main reason for their rejection in online banking.
- It was found that costs from the use of the Internet and from bank charges are of slight importance. This may be due to the fact that neither the charges nor the net-expenses are overpriced. Sixty - eight percent of the participants estimate it as a "non important" factor.
- On the other hand, the difficulties of the user's involvement in the electronic system are found to be of importance, as this restricts the control

that an individual exercises over the process (i.e. access to online services depends on access to the Internet). Only 27% consider the Internet's use as something extremely difficult while less than 20% believe that using electronic banking services is a wearisome task.

3.2.2 Willingness of use online banking services

A logistic regression model, was employed in order to identify the significance of the factors, which play a crucial role in an individual's decision whether or not to use online banking services and to estimate the probability of each individual using e-services. Logistic regression model estimates for each customer the logarithm of the probability of using on line services to the probability of not using online banking services.

Eight factorial variables are under examination for their contribution in the logit (p_i) of the i -th customer, which were x_{1i} = age, x_{2i} = education, x_{3i} = monthly income, x_{4i} = transactions' speed, x_{5i} = fear of change, x_{6i} = lack of Information, x_{7i} = difficulties of using Internet, x_{8i} = cost of using Internet own property.

The proposed logistic regression model is:

$$\begin{aligned} \text{logit}(p_i) &= \\ \log(p_i/(1-p_i)) &= \\ & 1.60 * x_{1,i} - 1.26 * x_{5,i} - 1.84 * x_{6,i} \\ & - 2.69 * x_{7,i} + 3.19 \end{aligned} \quad (13)$$

or

$$p_i = \frac{\exp(1.60 * x_{1,i} - 1.26 * x_{5,i} - 1.84 * x_{6,i} - 2.69 * x_{7,i} + 3.19)}{1 + \exp(1.60 * x_{1,i} - 1.26 * x_{5,i} - 1.84 * x_{6,i} - 2.69 * x_{7,i} + 3.19)}$$

The p_i represents the probability of the individual i to use on-line banking services, while the explanatory variables x_{1i} , x_{5i} , x_{6i} , and x_{7i} represent overall measures of factors regarding age, fear of change, lack of Information and difficulties of using Internet, respectively. Since our survey did not explicitly introduce cost and transactions' speed, the effects of these factors are parameterized in terms of the logistic constant.

The followings are depicted when we use the proposed generalized model

1. Individual's age is a significant determinant of someone's decision to use or not electronic services. This finding regards respect to all previous studies that defined the profile of online banking customer as a young man or woman who is familiar to technology advancements, and to PC and Internet

navigation. This result is consistent with hypothesis 5 (a).

2. The fear of change significantly affects the decision of use. An increase in this factorial coefficient would increase the willingness of use. This finding is consistent to hypothesis 3.
3. The factor "information" in someone's decision is of major importance. Hypothesis 6 is affirmed by this proposition.
4. User- friendliness is a significant determinant of an individual's acceptance or rejection of electronic services. This is consistent with hypothesis 2.
5. Variables education and monthly income are denoted as non-significant variables According to these results hypothesis 5(b) does not contribute to the willingness of using online banking services.
6. The results of variables transactions' speed, and Cost of using the Internet signifie that these variables could be significant if some of the parameters were to change (i.e. take into account and other variables). Hypotheses 1 and 4 seem to be less important in the willingness of use.
7. The constant coefficient β_7 is consistent with our construction, under which the transaction's speed and the cost of e-services are parameterized in the logistic constant. However as, the constant term is statistically significant, it is indicated that a rise in cost variable or in transaction's speed variable is relatively important with regard to decision of use.

4. Customer switching behaviour in greek banking services using survival analysis

4.1. A brief literature review

In nowadays, understanding and reacting to changes of customer behavior is an inevitable aspect of surviving in a competitive and mature market (Lariviere & Poel., 2004). Banks are facing the increased competition due to two different reasons: (a) the entrance of financial and insurance firms in the traditional banking market, and (b) the wide range of offered products and services to public. As a consequence the banking industry strives to succeed by putting the topic of rapid and changing customers needs to their agenda (Krishnan, Ramaswamy, Meyer & Damien, 1999).

The economic value of customer retention is widely recognized in the literature: (1) Successful customer retention lowers the need for seeking new and potentially risky customers and allows organizations to focus more accurately on the needs of their existing customers by building relationships

(Dawes & Swailes, 1999). (2) Long-term customers buy more and, if satisfied may provide new referrals through positive word-of-mouth for the company. (3) Long-term customers become less costly to serve due to the bank's greater knowledge of the existing customer and to decrease serving costs. (4) They tend to be less sensitive to comparative marketing activities (Ganesh *et al.*, 2000; Colgate *et al.*, 1996). (5) Losing customers not only leads to opportunity costs because the reduced sales, but also to an increased need for attracting new customers which is five to six times more expensive than customer retention (Athanasopoulos, 2000).

Many researchers in the literature have investigated the churn behavior of the banking customers.

Proportional hazard models are used by Bolton (1998) and Van den Poel & Lariviere (2004) in order to examine the link between customer satisfaction and retention. They found that (a) demographic characteristics, environmental changes and stimulating "interactive and continuous" relationships with customers, are of major concern when considering retention and that (b) customer behavior predictors only have a limited impact on attrition in terms of total products owned as well as inter-purchase time.

Regression models are used by Bloemer, Ruyter and Peeters, (1998) and Athanasopoulos (2000). They reveal that image is indirectly related to bank loyalty via perceived quality. They proposed an instrument of customer satisfaction in retail banking services. The empirical results have confirmed that customer satisfaction is a function of service quality (staff service and corporate image), price, convenience and innovation. Levesque & McDougall, (1996), point out that that service problems and the bank's service recovery ability have a major impact on customer satisfaction and intentions to switch. They identified the determinants, which included service quality dimension, service features, service problems, service recovery and products used. They concluded

The relationship between commercial banks and client companies was studied using t-test by Paulin *et al.*, (1998). Switching costs are increasingly finding their way into models of customer loyalty. According to Jones *et al.*, (2002), switching costs can be thought of as barriers that hold customers in service relationships. Krishnan *et al.*, (1999), via a Bayesian analysis, found that satisfaction with product offerings is a primary driver of overall customer satisfaction.

In summary, many studies investigated the problem of customers' switching behaviour. In this study we contribute in the existing literature at three different levels: (a) We examine the impact of qualitative groups of factors in retention behaviour (b) we examine customers' attrition considering the time aspect and (c) we use life tables in order to estimate the churn behaviour of clients in different periods of time.

4.2. Problem Description

Deregulation and increased competition from new products and delivery channels prompt banking industry to reconsider the economic value of customer retention. Customers' life cycles are becoming increasingly transitory due to severe impact of competitors' actions on existing relationships (Reinartz and Kumar, 2000). Typically, customers split their purchases among several competitive banks (Dwyer, 1997). In this study we examine the problem of customers' attrition, which is a classic problem of binary classification: churn behavior or not. More specifically we want to explain customers' switching behavior by examining a series of factors and by identifying the contribution of each one.

The study therefore, focused on determining the significance of the following factors:

H₁: There is no evidence that customers' gender affects their decision of breaking down their relationship with the bank.

H₂: Bank's credibility contributes positively in the duration of the relationship between customer and bank.

H₃: Customers' satisfaction is the most important factor for enlarging the duration of their cooperation with the bank

H₄: The quality of offered services and products determine the life cycle of the examined relationship.

H₅: "Interest Rate" is not recognized as a significant factor in customers' switching behavior

H₆: Individuals' educational level is considered as a valuable factor in the length of the examined relationship.

We use Life Tables to estimate the time of the churn event. A proportional hazard model is performed to find the factors which would pay in the reduction of customers' secession. Survival analysis is a collection of statistical methods for data analysis for which the outcome variable of interest is time until an event occurs (in our case: customer churn behavior) with the aim to develop

predictive models in which the risk of an event depends on covariates

We usually refer to the outcome variable as survival time, as it gives the time that an individual has survived over some follow-up period and we refer to the event as a failure. For some customers the time to failure (end the cooperation with the bank) may be observed completely, whereas for others we only have partial observation until some specific “censoring” time c . We denote by T , the random variable for a customer’s survival time.

The distribution of survival times is characterized by three functions:

- [1] The probability that a customer continues to cooperate with the bank longer than t , is defined by survival function $S(t)$

$$S(t) = P(\text{a customer continues to cooperate with the bank longer than } t) = P(T > t) \quad (14)$$

The cumulative distribution function of T is denoted by $F(t)$, and hence

$$S(t) = 1 - P(\text{a customer ends its relationship with the bank before time } t) = 1 - F(t) \quad (15)$$

- [2] The probability density function of the survival time, T , is denoted by $f(t)$ and is defined as the limit of the probability that a customer breaks down its relationship with the bank in the short interval $(t$ to $t + \Delta t)$ per unit width Δt .

It is expressed as (16)

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P\{\text{a customer ends its relationship with the bank in the interval } (t, t + \Delta t)\}}{\Delta t}$$

The density function is known as the unconditional failure rate.

- [3] The hazard function $h(t)$ of survival time T gives the conditional failure rate or the instantaneous failure rate (churn behavior). It is defined as the probability of failure during a very small time interval, assuming that the customer has decided to continue his/her cooperation with the bank to the beginning of the interval. It is expressed as (17)

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P\{\text{a customer of age } t \text{ ends its relationship with the bank in the interval } (t, t + \Delta t)\}}{\Delta t}$$

The hazard function is defined in terms of cumulative distribution function $F(t)$ and the probability density function $f(t)$ as

$$\begin{aligned} h(t) &= f(t) / \{1 - F(t)\} \\ &= f(t) / S(t) \end{aligned} \quad (18)$$

From equations (2) and (3) we conclude for the cumulative distribution functions

$$f(t) = \frac{d}{dt} [F(t)] = \frac{d}{dt} [1 - S(t)] \quad (19)$$

4.2.1 Life Table Analysis

The life table is a method for measuring churn behavior and describing survival experience of bank’s customers. These tables summarize the switching behavior of the customers for a specific period of time. Considering that the period under examination is divided to d intervals of width r_d , where $r_d = t_{d+1} - t_d$, provide information about the number of customers that enter in each interval, the number of customers exposed to risk of secession, the probability of churn behavior in the d -th interval per unit width and an estimation of the survival rate in this interval.

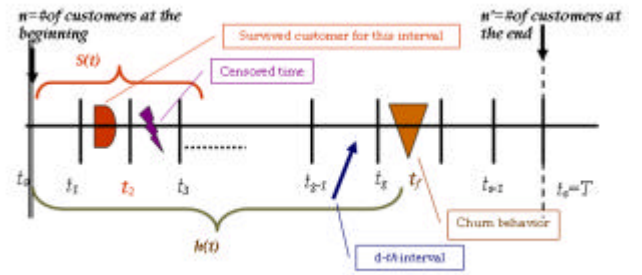


Figure 1 depicts the life cycle information in a period of time T . In the beginning of the study we observe the behavior of n customers, while at the end of the examined period the number has changed to n' as some clients broke down their cooperation with the bank (churn behavior) while for some others we don't have any information (censored behavior).

4.2.2 The Proportional Hazard Model

Let n be the number of customers and i their index. The Proportional Hazard Model gives an expression for the hazard at time t for the i -th customer with a given specification of a set of explanatory variables. We model these factorial variables, for each customer i , using the following p -dimension vector: $\underline{x}^i = [x_1^i, x_2^i, \dots, x_p^i]$. Each element x_j^i of \underline{x}^i represents products' attributes or demographical data for customer i . Note that j is the index of explanatory variables.

The Proportional Hazard Model proposed by Cox can be written as follows:

$$\begin{aligned} h^i(t, \underline{x}^i) &= h_0(t) e^{\left(\sum_{j=1}^p b_j x_j^i\right)} \\ &= h_0(t) \exp(b_1 x_1^i + b_2 x_2^i + \dots + b_p x_p^i) \end{aligned} \quad (20)$$

In (20) $h^i(t, \underline{x}^i)$ represents the hazard for the i -th customer at time t , $h_0(t)$: represents the baseline hazard function and b_j ; represents the coefficient of the covariate j , x_{ij} represents the value of the i -th customer for the varying covariate j .

4.3. Results from Survival Analysis

We have considered 6 independent x_j^i ($i=1, \dots, n$) and ($j=1, \dots, 6$) factorial variables which are x_{1i} = gender, x_{2i} = bank's accuracy, x_{3i} = customers' satisfaction, x_{4i} = service quality, x_{5i} = interest rates, x_{6i} = educational level. A European financial services company that offers banking and insurance services provided the data for this study.

Investigating the survival probabilities, due to life tables we conclude that churn behavior differs among the range of the factors. Some results from our analysis are presented below:

- The median time, is estimated equal to 45,6 years old for men and to 45,08 for women. In order to examine the power of the Hypothesis H1 we used Wilcoxon test, according to which we found that there is no difference between men and women in the duration of their relationship with the bank.
- A binary variable also represents the factor "bank's credibility". The median time of survival for the customers who consider "bank's credibility" as a strong factor for continuing their cooperation with bank is estimated up to 38.67 years old and up to 45.65 for the customers belonging to second category. This finding is consistent with Hypothesis H2.
- "Customer Satisfaction" is modelled by an ordinal variable with 7 scales. (1= totally – satisfied, 7= no-satisfied). The median time for the fully satisfied customers is calculated equal to 43.13 years old, while the respectively value for the dissatisfied customers is 38 years. Hypothesis H3 is affirmed by this proposition.
- "Service quality" is also model by an ordinal variable with 7 scales from 1 (=non important) to 7 (=extremely important). The majority of the customers of our samples rates this factor with values between 4 to 7. The median life for these categories is estimated 44.06 years old. These results are consistent with Hypothesis 4.
- The deregulation and the reformation of economic markets turned interest rates to a very important issue for the customers. Variable "interest rates" is modelled by a seven-item scale. The likelihood of switching behaviour for the customers that characterized the interest rates as low or high is higher than the customers who rank interest rates with values 1-2, 4 and 6-7. These proportions suggest that the hypothesis H5 does not totally approved.
- The variable "Educational level" has four categories (1=primary school or less, 2= high

school, 3= university studies and 4=postgraduate studies). According to the results, apart from the customers with low educational level which seem to break down the relationship with the bank in the age of 45-55 years old, the impact of this variable is increasing the duration of the examined relationship for the other three categories is limited. Thus, the power of Hypothesis 6 isn't strong enough.

Churn Behavior and Churn Predictors using Cox Regression

Figure 2 presents the survival function at mean of covariates for all customers included in this study. The plot shows that the expected decreasing shape suggesting the fact that the longer an individual has been a bank customer the smaller the probability of survival. From the graph we note that customers experience a high switching probability (>0.6) (switching probability = $1 - \text{survival probability}$ according to equation (15)) in the age of 35-45. After the age of 50 years old the probability of survive decreases.

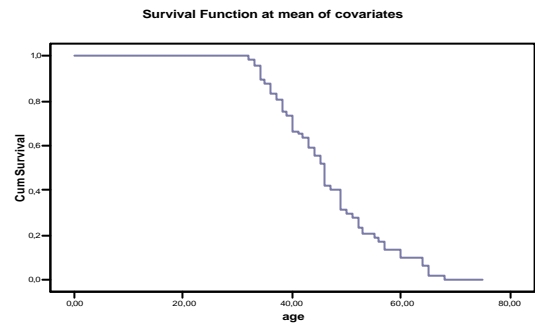


Figure 2: Survival Behavior for all customers

The impact of the observed covariates on the retention is investigated by the use of the Cox Proportional Model, as we described it in equation (16)-(20).

A number of interesting findings emerge from proportional hazard model:

- In terms of demographic characteristics, the lower educational level has a positive impact in the limitation of the risk of churn behavior
- Both the quality of the offered services and the bank's brand name has a negative effect in the decrease of the attrition. We conclude to the same argument for customers' satisfaction
- The values of beta coefficients (b_j) for interest rates signify that "high" and "low" interest rates increase the risk of retention.

5. Conclusions

In this study an integrated approach for reconfiguring a community branch network according to the dictates of the market is developed, the internal bank resources and the strategic policy constraints. a linear programming model which was solved via an iterative algorithm that targets the optimum number of branches and the optimum mix of services that each branch should offer, is proposed. The results indicate that through the this approach, the branch network at the community level can be streamlined and transformed into an effective, revenue-generating group of bank nodes.

Secondly a framework in predicting consumer's behavior in new products and services is proposed. An estimated probability of an individual to use or not to use online banking services is provided by a generalized linear model. Individual's age, the difficulties of using the Internet, the fear of the changes provoked to the banking sector due to technological development and the lack of information concerning products and services provided to customers through electronic delivery channels, are identified as crucial factors that affect someone's decision to use or not to use online services.

Finally the churn behavior of bank customers is investigated in greek market. With the help of Life Tables and the proportional hazard model we find out that the quality of the offered banking products and services in combination with the bank's brand name have a positively effect in the decrease of switching behavior while demographical characteristics such as gender and educational level have a limited impact.

Considering this findings in the context of customer behavior we conclude that a bank needs to develop a CRM systems, in order to estimate the adoption of new products and services, decrease the rate of retention and brings on its performance. Opportunities for further research reside in determining segments of customers with common characteristics that will contribute in the above three objectives.

References

Anderson, E., Fornell, C., & Lehmann, D. (1994). Customer Satisfaction, Market Share, and Profitability: Findings from Sweden. *Journal of Marketing*, **58**, 53-66.

Athanassopoulos, A.D, & Curram, S. (1996). A comparison of Data Envelopment Analysis and Artificial Neural Networks as Tools for Assessing the Efficiency of Decision Making Units. *Journal of the Operational Research Society* **47**, 1000-1016

Athanassopoulos, D.A. (2000). Customer Satisfaction Cues to Support market segmentation and Explain Switching Behavior. *Journal of Business research*, **47**, 191-207.

Avkiran, N.K. (1997). Models of retail performance for bank branches: predicting the level of key business drivers. *International Journal of Bank Marketing*, **15**(6), 224-237.

Avkiran, N.K. (1999). An application reference for data envelopment analysis in branch banking: helping the novice researcher. *International Journal of Bank Marketing* **17**(5), 206-220.

Berry, L. (1995). Relationship Marketing of Services-Growing Interest, Emerging Perspectives. *Journal of the Academy of Marketing Science*, **23**(4), 236-245.

Bloemer, J., Ko de Ruyter, & Peeters, P. (1998). Investigating drivers of bank loyalty: the complex relationship between image, service quality and satisfaction. *International Journal of bank Marketing*, **16**(7), 276-286.

Bolton, N, R. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: the role of satisfaction. *Marketing Science*, **17** (1), 45-65.

Boufounou, P.V.(1995). Evaluating bank Branch Location and Performance: A case study. *European Journal of Operational Research* **87**, 389-402.

Bomil, S., Ingo, H. (2002) Effect of trust on customer acceptance of Internet banking. *Electronic Commerce Research and Applications*, **247-263**

Burton, D. (1990). Competition in the UK retail financial service sector: some implications for the spatial distribution and function of bank branches. *Service Industries Journal*, **10**(3), 571-588.

Camanho A.S., & Dyson, R.G. (1999). Efficiency, Size, Benchmarks and Targets for Bank branches: an application of data envelopment analysis. *Journal of the Operational Research Society* **50**, 903-915.

Clawson, J. (1974). Fitting Branch locations, performance standards, and marketing strategies to local conditions. *Journal of Marketing* **38**(1), 8-14.

Colgate, M., Stewart, K., & Kinsella, R. (1996). Customer defection: A study of student

- market in Ireland. *International Journal of Bank Marketing*, 14(3), 23-29.
- Dawes, J., & Swailes, S. (1999). Retention sans frontiers: Issues for financial service retailers. *International Journal of Bank Marketing*, 17(1), 36-43.
- Delvin, J. (1995). Technology and innovation in retail banking distribution", *International Journal of Bank Marketing*, 13(4), 19-25.
- Doyle, P, Fenwick, L. & Savage, G.P. (1979). Management planning and control in multi-branch banking. *Journal of Operational Research Society* 30(2), 105-111.
- Drake, L., & Howcroft B. (1994). Relative efficiency in the branch network of a UK Bank: An empirical study. *Omega* 22(1), 83-90.
- Frei, X.F. & Harker, P.T. (1999) Measuring the Efficiency of Service Delivery Processes: An Application to Retail Banking. *Journal of services Research 1*, ; 300-312.
- Ganesh, J., Arnold, J., & Reynolds, K.E. (2000). Understanding the customer base of service providers: An examination of the differences between switchers and stayers. *Journal of Marketing*, 64(3), 65-87.
- Garvin, D. (1988). Managing quality: The strategic and competitive edge. *The Free Press*, New York.
- Hitt, L., Frei, F. (2002). Do Better Customers Utilize Electronic Distribution Channels? The Case of PC Banking, *Management Science*, 48 (6), 732-748.
- Hansen, L. (1990). Measuring Branch Performance. *Credit Union Management* 13(3), 8-11
- Heald, G.I. (1972). The application of the automatic interaction detector (A.I.D.) programme and multiple regression techniques to the assessment of store performance and site selection. *Operational Research Quarterly* 23(4), 445-57.
- Hempel, G. & Simonson, D. (5th Edition) (1998). *Bank Management*. John Willey& Sons, Inc.
- Jones, A.M., Mothersbaugh, L.D., & Beatty, E.S. (2002). Why customers stay: measuring the underlying dimensions of services switching costs and managing their differential strategic outcomes. *Journal of Business Research*, 55, 441-450.
- Jun, M., & Cai, S. (2001). The key determinants of Internet Banking service quality: a content analysis. *International Journal of Bank Marketing*, 19(7), 276-291.
- Karjaluoto,?, Mattila, ? , & Pento, T. (2002). Factors underlying attitude formation towards online banking in Finland. *International Journal of Bank Marketing*, 20(6), 261-272.
- Kennington, C., Hill, J., & Rakowska, A. (1996). Consumer selection criteria for banks in Poland. *International Journal of Bank Marketing*, 14(4), 12-21
- Koufaris, M., Kambil, A., & LaBarbera, A.P.(2002). Customer behaviour in web based commerce: an empirical study. *International Journal of Electronic Commerce*, 6(2), 115-138.
- Krishnan, M.S., Ramaswamy, V., Meyer, C.M, & Damien, P. (1999). Customer Satisfaction for Financial Services: the role of products, services, and Information Technology. *Management Science*, 45(9), 1194-1209.
- Lariviere, B., & Van den Poel, D. (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modelling: the case of financial services. *Expert Systems with Applications*, 27, 277-285.
- Levesque, T., & McDougall, H.G.G. (1996). Determinants of customer satisfaction in retail banking, *International Journal of Bank Marketing*, 16 (1), 24-31.
- Liao, Z., & Cheung, M. (2002). Internet-based e-banking and consumer attitudes: an empirical study. *Information and Management*, 39, 283-295.
- Luarn, P., & Lin, H. (2005). Toward an understanding of the behavioral intention to use mobile banking. *Computers in Human Behavior*, 21, 873-891.
- Mols, N. (1999). The Internet and the banks' strategic distribution channel decisions. *International Journal of Bank Marketing*, 17(6), 295-300.
- Mountinho, L., Davies, F., Deng, S., Peris, S., & Alcaniz, J. (1997). *International Journal of Bank Marketing*, 15(3), 99-105.

Oral, M., Kettani, O., & Yolalan, R. (1992). An empirical study of analyzing the productivity of bank branches. *IIE Transactions* 24, 166-176.

Paulin, M., Perriien, J., Ferguson, J.R., Salazar, A.M.A., & Seruya, M.L. (1998). *Relational norms and client retention: external effectiveness of commercial banking in Canada and Mexico*”,

Schaffnit C, Rosen D., & Paradi J.C.(1997). Best practice analysis of bank branches: An application of DEA in a large Canadian bank. *European Journal of Operational Research* 98, 269-289.

Smith, R.L. & Schweikart, J.A.(1992). A better way to assess branch profitability. *ABA Banking Journal* 84(4), 54-56.

Thygeson, K.J. (1991). Modeling branch profitability. *Journal of Retail Banking* 13(3), 19-24.

Vassiloglou, M. and Giokas, D. (1990). A study of the relative efficiency of bank branches: An application of data envelopment analysis. *Journal of the Operational Research Society* 41(7), 591-597.

Wang, Yu-S., Wang, Yu-M., Lin, H.H., & Tang, T.I. (2003). Determinants of user acceptance of Internet Banking: an empirical study“, *International Journal of Service Industry Management*, 14(5), 501-519

Zineldin, M. (1996). Bank strategic positioning and some determinants of bank selection”, *International Journal of Bank Marketing*, 14(6), 12-22.