Quaderni di Dipartimento

Monitoring and Improving Greek Banking Services Using Bayesian Networks: an Analysis of Mystery Shopping Data

Claudia Tarantola
(Università di Pavia)

Paola Vicard
(Università Roma Tre)

Ioannis Ntzoufras
(Athens University of Economics and Business)

160 (01-12)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Gennaio 2012

Monitoring and Improving Greek Banking Services Using Bayesian Networks: an Analysis of Mystery Shopping Data

Claudia Tarantola, Paola Vicard and Ioannis Ntzoufras*

Abstract

Mystery shopping is a well known marketing technique used by companies and marketing analysts to measure quality of service, and gather information about products and services. In this article, we analyse data from mystery shopping surveys via Bayesian networks in order to examine and evaluate the quality of service offered by the loan departments of Greek banks. We use mystery shopping visits to collect information about loan products and services and, by this way, evaluate the customer satisfaction and plan improvement strategies that will assist Banks to reach their internal standards. Bayesian Networks not only provide a pictorial representation of the dependence structure between the characteristics of interest but also allow to evaluate, interpret and understand the effects of possible improvement strategies.

Keywords: Bayesian networks, Customer satisfaction, Mystery shopping, Service quality improvement.

*Claudia Tarantola is an Assistant Professor at the Department of Economics and Quantitative Methods, University of Pavia, claudia.tarantola@unipv.it; Paola Vicard is an Associate Professor at the Department of Economics, University Roma Tre, vicard@uniroma3.it; Ioannis Ntzoufras is an Associate Professor at the Department of Statistics, Athens University of Economics and Business, ntzoufras@aueb.gr

1 Introduction

The banking industry is a highly competitive and customer oriented organisation. Customer retention and attraction is a core element of its managing strategy; customer service is one of the factors allowing to differentiate a bank from its competitors.

Roughly speaking, customer satisfaction refers to the extent to which products and services supplied by a company meet or exceed customer expectation. Customer satisfaction levels can be measured using survey techniques and evaluation questionnaires. High levels of customer satisfaction indicate a good performance of the business since satisfied customers are most likely to be loyal to the specific company and use a wide range of services. Understanding which elements influence customer satisfaction is important not only to describe the actual situation but also to plan and implement possible improvement actions.

In this paper we use Bayesian Networks (BN hereafter) to analyse data gathered from mystery shoppers' report. To our knowledge, this is the first time that these techniques are used in combination. We present a real data analysis concerning customer evaluation of service provided by the loan unit of Greek banks. For some recent works regarding customer satisfaction analysis of Greek banks see e.g. Mihelis *et. al* (2001), Grigoroudis *et al* (2002), Mylonakis (2009) and Kagara and Voyiatzis (2010).

Mystery shopping is a well known marketing technique used by companies and marketing analysts to measure quality of service, and gather specific information about products and services. Nowadays, it is one of the most used techniques for performance evaluation of banks; see e.g. Schrader (2006), Sherman and Zhu (2006), Roberts and Campbell (2007) and references therein.

A BN is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a Directed Acyclic Graph (DAG). The use of a graph, as a pictorial representation of the problem at hand, simplifies model interpretation, and facilitates communication and interaction among experts with different backgrounds. For these reasons, BNs are widely applied in different fields for the analysis of multivariate data, see Neapolitan, 2004.

Recently BNs have been successfully applied to the analysis of customer satisfaction data, see for example Salini and Kennet (2007) and Renzi *et. al* (2009). Providing a DAG representation for the problem under investigation, BNs allow to easily identify the key elements influencing customer satisfaction. Furthermore they can be used to simulate improvement strategies, getting reliable results in a straightforward manner.

The paper is organised as follows. In Section 2.1 we present the mystery shopping methodology. BNs together with the procedure to construct them are illustrated in Sections 2.2 and 2.3. Section 3 is devoted to the application of BNs to service quality improvement in Greek banks. Finally, in Section 4 we end up with some comments and final remarks.

2 Background and Preliminaries

2.1 Mystery Shopping

Mystery shopping is a well established methodology which was introduced in the early 1940s primarily by the management of banks and retail chain stores to assess the integrity of their employees (Zikmund *et al.*, 2009). Nowadays there are hundreds of companies providing services related to mystery shopping surveys (see for example `http://www.mysteryshop.org/`). For a comprehensive introduction and an exhaustive discussion on the topic we refer the reader to the publications of Wilson (2001) and Saha (2009).

Mystery shoppers are well trained individuals used to anonymously evaluate and monitor customer satisfaction and quality of the service in different sectors (*i.e.* top retail outlets, restaurants, cinemas, banks, theatres, travel companies, hotels, spas, cruise companies, airlines, amusement parks and leisure organizations). They act as normal or potential customers and make unannounced visits to the company. Through the use of mystery shoppers it is possible to monitor, not only the quality of the service, but also the efficiency of the process and the procedure followed to deliver the service. After each visit the mystery shoppers complete a predefined questionnaire report concerning their service experience. This report usually includes numerical ranking on a series of statements, check-lists, and open-ended questions regarding the shopper general impression. The results provided by such surveys can be used to evaluate and compare the performance of different companies or branches of the same company or individual employees. This information also assists company managers to monitor how the performance changes over time and to identify areas that require improvement.

Nowadays mystery shopping is one of the most popular techniques used to evaluate customer satisfaction in the banking sector. Banking mystery shoppers provide information about the quality of financial products, the efficiency of the bureaucratic procedures and the politeness of the employees, amongst others. The long-term aim of their visit is to identify areas that require improvement and, by this way, offer advice and information concerning managerial actions that will improve the overall profile of the company. To be more specific, banking mystery shoppers assess the administrative functions and interpersonal skills of bank employees. They evaluate sales effectiveness of platform and teller staff by analysing whether the bank staff listens to its customers, how friendly tellers are, how much time the customers take to get to a teller and many other aspects.

2.2 Bayesian Networks

A BN is a multivariate statistical model that uses a DAG to represent statistical dependencies among variables; see for example Jensen (1996) and Cowell *et al.* (1999). It combines features of graph and probability theory. The term "Bayesian" does not refer to the Bayesian inferential paradigm, but is due to an efficient information propagation algorithm based on the Bayes theorem. More precisely, a BN is characterised by:

(i) a DAG showing the set of dependencies among variables and

(ii) an inferential engine to make inference on the parameters of the model.

Here we summarize the main elements and terminology about BNs that will be used in this paper. A DAG is a mathematical object defined by the pair $D = (V; E)$, where V is the set of nodes and E is the set of directed edges (arrows) connecting pair of nodes. Nodes represent the variables of interest of the studied problem. Here, each node is associated with a random variable (item of the questionnaire) relative to one of the characteristics of the Bank surveyed and reported by the mystery shopper. Arrows represent relations (or, more statistically speaking, dependencies) between variables. The directed graph is said acyclic (DAG) since cycles are forbidden, *i.e.* it is not possible to start from a variable (node) and, following the directions of the arrows/edges, go back to the starting node. An example of DAG is provided in Figure 1. If an arrow points from variable X_i to variable X_j then X_i is called parent of X_j, denoted by $pa(X_j)$, and the two variables are dependent. For example in Figure 1 variables X_1 and X_5 are parents of X_4.

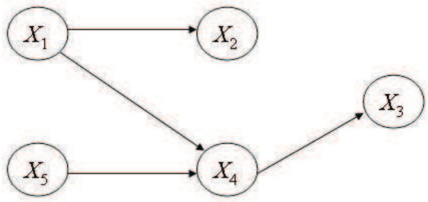

Figure 1: An example of DAG

The graph can be used to describe and read marginal and conditional independencies among the variables under consideration. The absence of an edge between two variables might indicate a conditional independence statement. Let us consider variables X_1 and X_3 in Figure 1. In order to go from node X_1 to node X_3 it is necessary to pass by node X_4. This allows us to say that variables X_1 and X_3 are independent conditionally on variable X_4, and therefore, knowing the observed value of variable X_4, variable X_1 becomes uninformative for X_3. In other words, suppose X_3 denotes the overall satisfaction, X_4 has a direct effect while X_1 an indirect effect on it. Therefore if an intervention on X_4 is done, intervening on X_1 has no additional effect on the overall satisfaction (X_3). In general any graph configuration such as $X_i \to X_j \to X_k$, indicates that variables X_i and X_k are independent given X_j. Consider now variables X_2 and X_4 in Figure 1; they are separated by node X_1 and therefore they are independent conditionally on variable X_1. In general for a graph configuration such as $X_i \leftarrow X_j \rightarrow X_k$, we have that variables X_i and X_k are independent given X_j. This means that if two quality aspects, say X_i and X_k have a common parent X_j, then an intervention on X_j makes X_i and X_k independent.

Differently from the previous two cases, a graph configuration such as $X_i \rightarrow X_j \leftarrow X_k$ indicates that variables X_i and X_k are dependent given X_j. For example, if we consider nodes X_1 and X_5 in Figure 1, they are not independent given node X_4, that is having information about variable X_4 makes its two parents, *i.e.* X_1 and X_5, dependent. For a detailed and rigorous account on the methods used to read independencies from a DAG we refer to Lauritzen (1996).

Given a network for K variables (*i.e.* the set of nodes V is made up of K nodes), each node X_i, $i \in V$, is associated with the conditional probability distribution of the corresponding variable of interest given its parents, $p\big(X_i|pa(X_i)\big)$. We can factorise the joint probability distribution of $(X_1,\ldots,X_K)$ according to the dependencies encoded in the DAG

$$p(X_1,\ldots,X_K) = \prod_{i\in V} p\Big(X_i|pa(X_i)\Big). \tag{1}$$

If a node, say X_i, has no parents then it is associated with the marginal distribution of X_i, $p(X_i)$. For the DAG in Figure 1 we have:

$$p(X_1,X_2,X_3,X_4,X_5) = p(X_1)p(X_2|X_1)p(X_3|X_4)p(X_4|X_1,X_5)p(X_5).$$

In order to use BNs, the graphical structure must be first specified and the corresponding parameters of the distributions in (1) must be estimated. Once the model structure and parameters have been estimated, efficient information propagation techniques can be exploited in order to assess the effect of changes in the status of specific variables on the behaviour of the remaining variables; see Jensen (1996) and Cowell, *et al.* (2001).

2.3 Structural Learning for BNs

If the structure of the phenomenon under consideration is known, the DAG can be built manually on the basis of expert knowledge; if the subject-matter knowledge is not strong enough to draw the graph, the network structure is necessarily learnt (estimated) from the available data.

Constructing the network by using expert knowledge is a cumbersome task, hence the dependence structure is usually learnt directly from data. The algorithms to learn the network structure from data can be grouped in two principal categories: score-based algorithms and constraint-based algorithms; for more details see Cooper and Herskovits (1992) and Neapolitan (2004).

The score-based algorithms explore the space of all possible DAGs in order to find the one that minimises a given score function. On the other hand, constraint-based algorithms carry out a series of independence tests and construct a graph which satisfies the discovered independence statements.

For our application we use the software Hugin (www.hugin.com), and we apply a two step procedure (described in Section 3.2) that combines characteristics of both algorithms.

3 Service Evaluation and Improvement in Greek Banks

3.1 Data Description

The mystery shopping data were collected by a female master student of the University of Aegean (Sergianiti, 2003). The aim of the research was to evaluate and compare the quality of the services offered by loan units of five popular banks operating in Greece in 2003. The names of the banks are suppressed for privacy reasons but they are available in Sergianiti (2003). The data consist of 128 mystery shopping reports conducted from late April to late August 2003 in five different cities: Athens, Ioannina, Prevesa, Arta and Chios.

The mystery shopper behaved like a normal customer, examining the internal environment and the approach of the bank loan consultant and interviewing the consultant about "low consumer loans" (up to 3000€). After each visit she filled a report form including checklist items, rating scales, and additional open-ended comments, see Sergianiti (2003) for details. The report consists of 45 items (variables) in total, but only 19 variables present a significant association being directly or indirectly connected to the main response variables of bank evaluation. Most variables were measured on a 5 level scale where 1 denotes the worst judgement and 5 the best one. In all cases the level 1 was never o very rarely chosen so that levels 1 and 2 were merged to avoid problems connected to highly sparse tables when learning the network by constraint-based algorithms (since they rely on chi-square type independence tests). For some variables (“Waiting time”and “Office cleanliness”) even level 2 was never o very rarely used in the reports. In the Appendix a short description of the examined variables is provided.

3.2 Learning the Network for Greek Banks

The network modelling the dependence structure among the customer satisfaction variables has been learnt by means of the following two-step procedure:

STEP 1: We created an initial draft of the network by means of the Chow-Liu algorithm, Chow and Liu (1968). This was done to solve sparsity problems. Moreover logical constraints, such as target variables having no outgoing arrows towards any of the other variables, have been taken into account. A similar idea has been also applied by Cheng *et al.* (2002).

STEP 2: We run the NPC (Necessary Path Condition) algorithm, Steck (2001), using as a set of constraints the tree selected via the Chow Liu procedure and the logical ordering among the variables. Using the Chow-Liu algorithm as a preliminary step of our procedure, we reduced the dimensionality of the space to explore with the NPC algorithm. Furthermore, the NPC algorithm allows the user to choose among independence equivalent models the one most suitable for the analysed problem.

In order to complete the construction of our model, we need to estimate the conditional distributions from the data; this is achieved via the EM algorithm whose version for BNs has been proposed by Lauritzen (1995).

3.3 Description of the selected network

The network in Figure 2 represents the selected dependence structure.

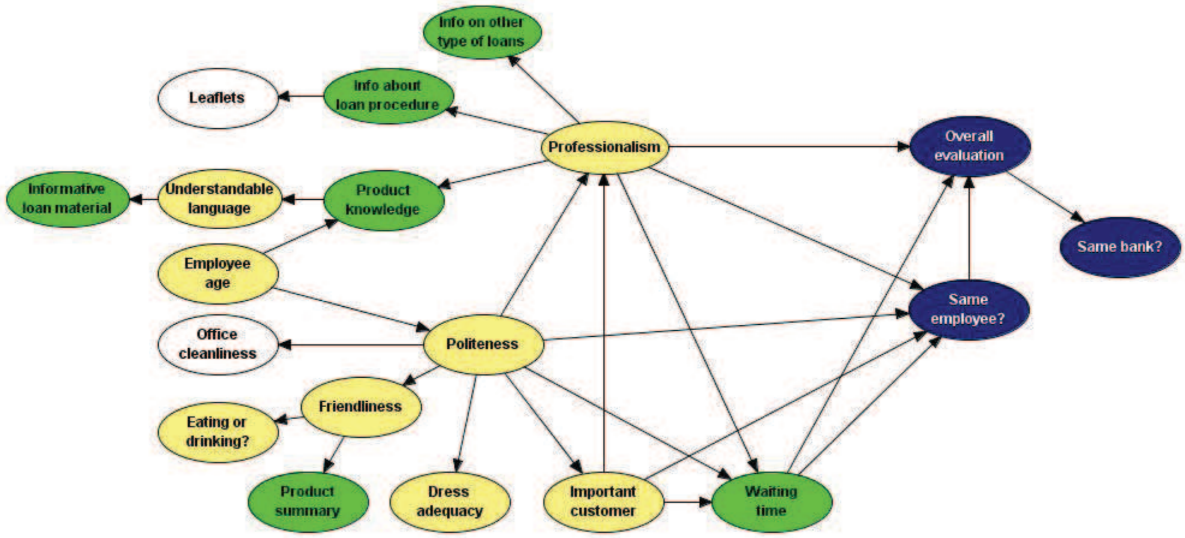

Figure 2: Selected network structure for mystery shopping data

Black coloured nodes are those directly related to the evaluation of the service (our target variables), describing whether the customer wishes to be served again by the same bank or employee ("Same bank?" and "Same employee?") and the overall score ("Overall evaluation"). Grey coloured nodes refer to service organization aspects and employee items while white nodes to bank items. Moving through the network from one node to another (not necessarily according to the arrow directions) we can identify which variables influence directly or indirectly our target ones. For example "Overall evaluation" is directly affected by the items "Professionalism" and "Waiting time", and it is indirectly influenced by many variables, some of which very far in the graph such as "Informative loan material" and "Product summary". Knowing which variables have a direct or indirect influence on the main response variables and identifying the path (sequence of nodes connected by arrows) linking a variable to a target node, is of relevant importance in order to plan and develop improvement strategies.

Figure 3 displays the marginal probability tables estimated from the data. The node tables show, for each variable, the estimated probabilities that banks got the different possible rates. This information gives a picture of the starting point before simulating and eventually implementing improvement actions.

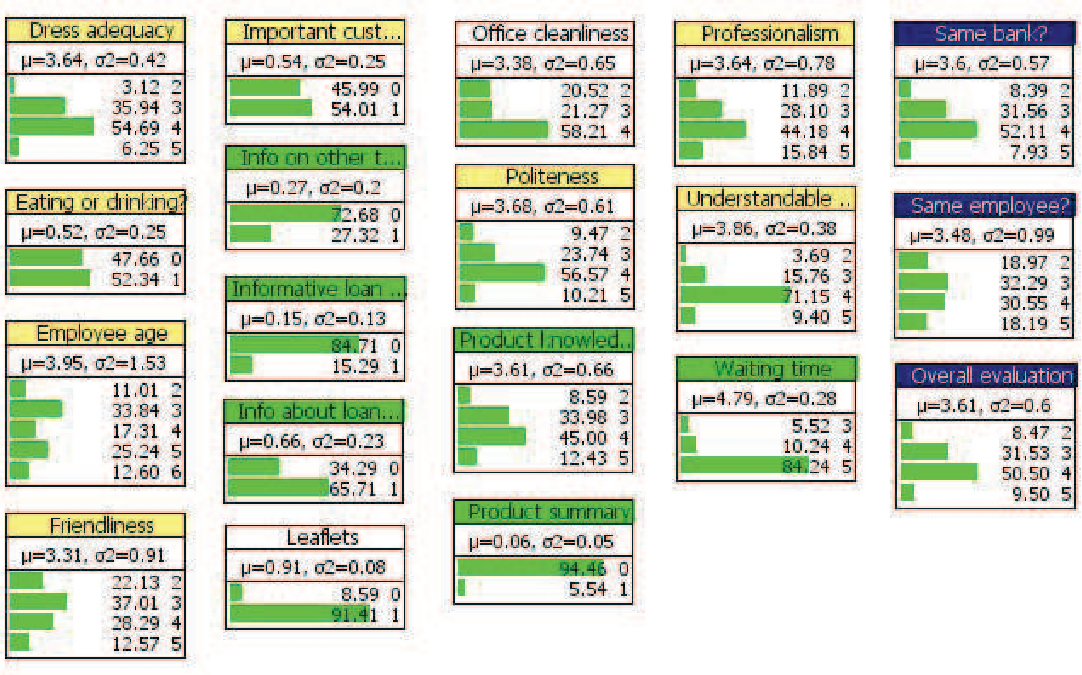

Figure 3: Estimated marginal probabilities (percentage values) for the customer satisfaction network in Figure 2

3.4 Using the Network – Examination of Different Scenarios

Once the model has been estimated, we can address a number of queries about the effect of strategic actions that will improve the overall evaluation and picture of the company. Different possible scenarios can be examined by inserting and propagating the appropriate evidence throughout the network. The effect of a strategic improvement action on a specific variable can be obtained in a mouse-click time by the evidence propagation algorithm. For example, if we wish to examine the effect of improving the professionalism of the employee, we might assign probability 100% to state 5 of "Professionalism" and then we immediately obtain the effect of this change on all the remaining variables of the network, *i.e.* the updated marginal probability tables of all variables. Assigning probability 100% to the highest value of a given variable allows to know the improvement margins of a possible action.

In the following, we illustrate different improvement scenarios.

Scenario A: Improving Zero Cost Variables

We study the effect of an improvement action on the variables “Product summary” and “Informative loan material” that can be handled with limited or zero cost. We simulate the realistic situation in which all the employees summarise the proposed products and give printed material regarding the loan procedure at the end of the meeting with the customers. “Product summary” and “Informative loan material” are two binary variables whose state 1 means that a product summary is done and informative material is given. We can insert this evidence in the network by double clicking on state 1 of both variables, see Figure 4. We note that even if these variables are not directly related to the main response variables of the bank evaluation, the effects of these improvement actions are substantial. The percentage of the highest evaluation score of variable “Same employee?” rises from 18.19% to 62.71% while the corresponding percentage of visiting the same bank branch (score 4 or 5) rises from 60.04% to 84.95%. Moreover, the percentage of the highest “Overall evaluation” score rises from 9.5% to 33.6%.

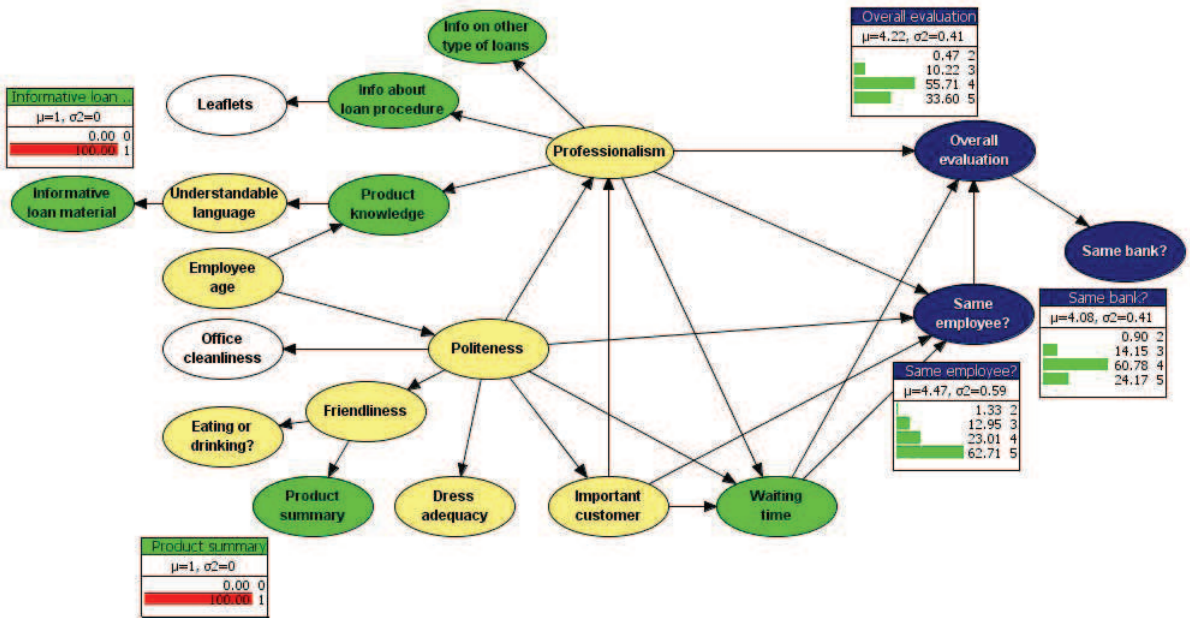

Figure 4: Effect of an improvement action on “Product summary” and “Informative loan material” on the global evaluation of the service

Scenario B: The Effect of Professionalism

Examining the graph in Figure 2 we notice the central role played by variable “Professionalism”. In fact, this variable is both directly and indirectly related to the target ones. Even more importantly, professionalism is directly logically connected to face-to-face service. Professional employees enhance the image of the bank, deal well with customers, and help the bank grow and succeed.

We first simulate the optimistic situation in which all the employees present the highest level of professionalism. This improving action has a tremendous and immediate effect on all the three

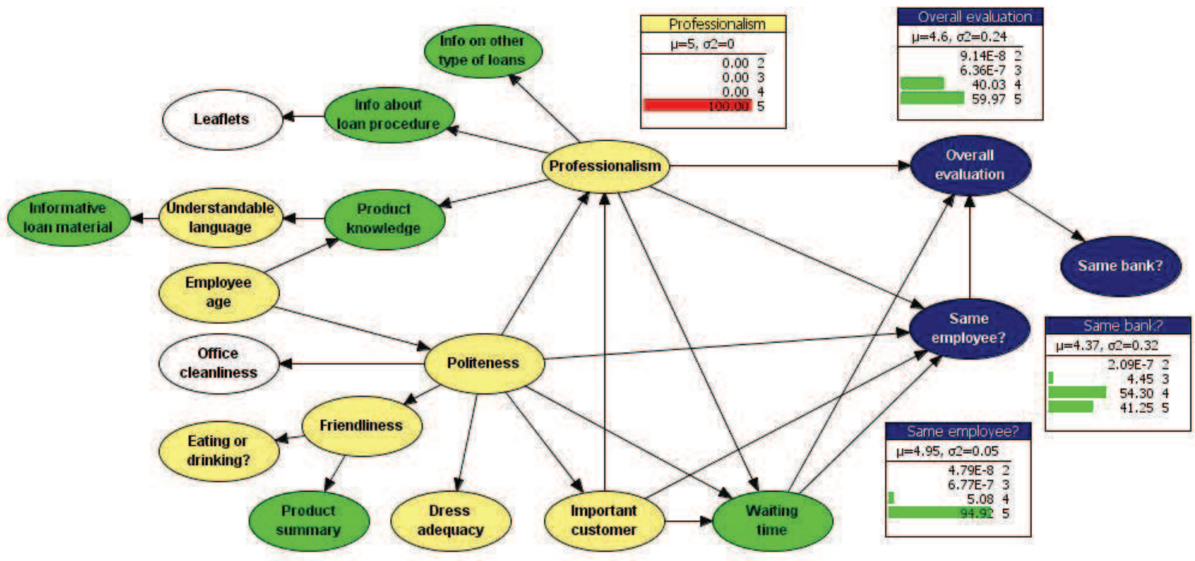

Figure 5: Effect of an improvement action on “Professionalism” on the global evaluation of the service

target variables, see Figure 5 . In fact for “Overall evaluation”, “Same employee?” and “Same bank” the percentage corresponding to the highest evaluation is now approximatively equal to 59.97%, 94.92% and 41.25% respectively, whereas initially it was 9.5%, 18.19% and 7.93%.

We then simulate an intermediate and more realistic target situation where the evaluations 2 and 3 for professionalism are not chosen. The effect of this milder improvement action is still remarkable as shown in Figure 6.

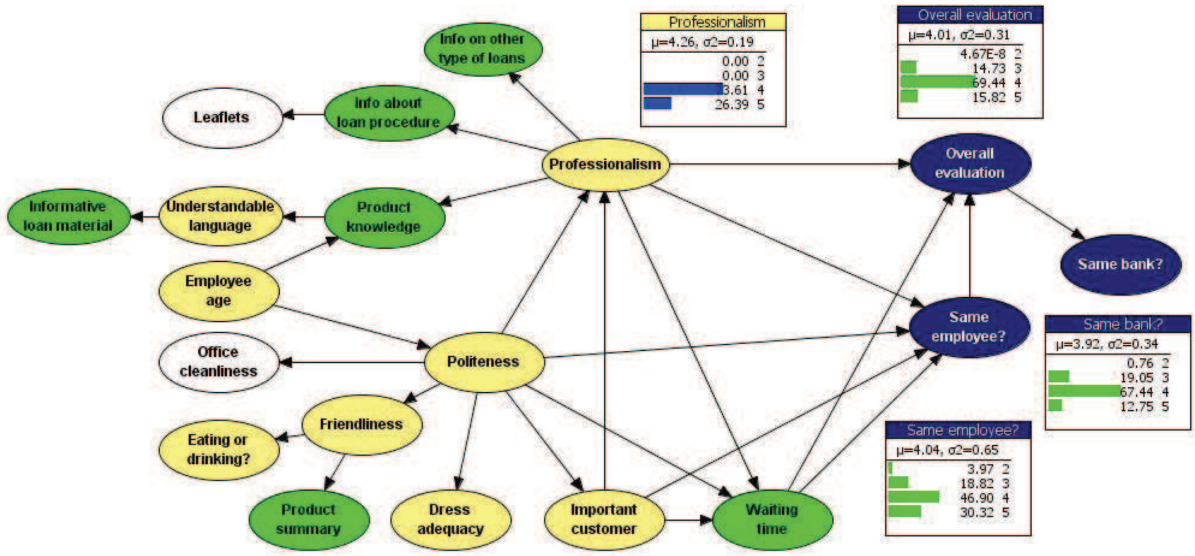

Figure 6: Effect of avoiding levels 2 and 3 of variable "Professionalism" on the global evaluation of the service

Scenario C: Improving Characteristics Related to Professionalism

"Professionalism" is not a tangible and measurable characteristic of employees, and a direct intervention on it may not be straightforward. Hence, we considered the effect of improvement actions on more tangible variables directly and/or indirectly related to it. We study the effect of an intervention on logical neighbours of variable "Professionalism", that is "Product knowledge" , "Information about loan procedure" and "Information on other type of loans".

We compare the effects of different improvement strategies (for sake of parsimony we do not report here the results) and we discover that a key role is played by variable "Product knowledge". Hence, we simulate the situation in which all the employees have an excellent knowledge of the product they are selling; see Figure 7. This action leads to an increase of the probabilities of scores 4 and 5 of 'Professionalism", consequently improving the overall satisfaction. More precisely, for the variable "Same employee?" the probability of the highest score rises from 18.19% to 76.88%, while that of scores 4 and 5 summed together rises from about 48.74% to 93.5%. Finally, notice that for the variable "Overall evaluation", the top score probability rises from 9.5% to 47.26%.

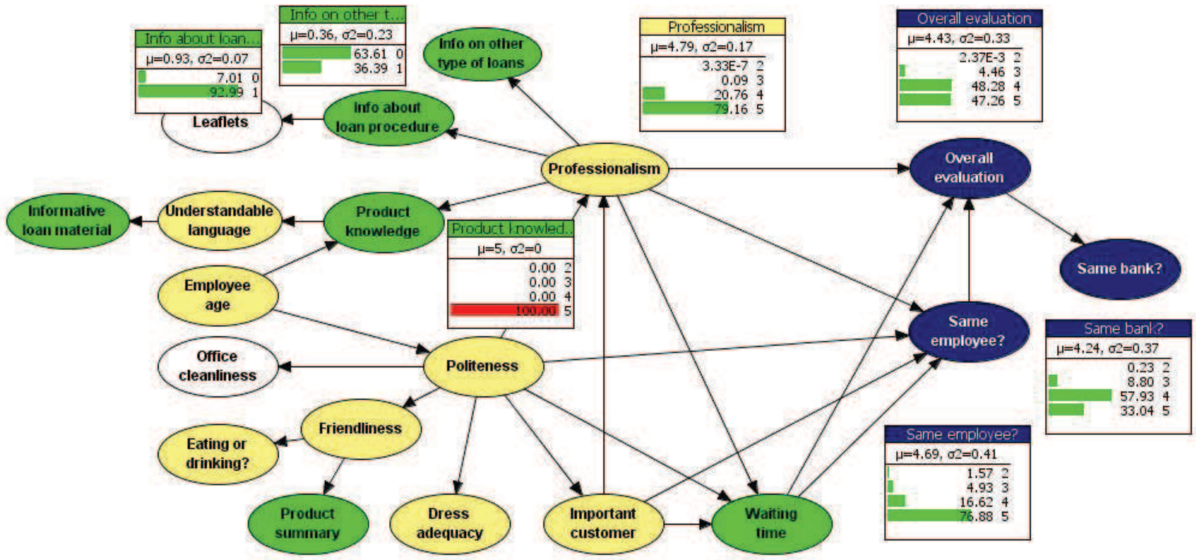

Figure 7: Effect of an improvement action on "Product knowledge" on variable "Professionalism", and on the global evaluation of the service

Scenario D: The Importance of Spoken Communication

In this scenario we study the effect of variable "Understandable language", indirectly related to "Professionalism" and to the response variables of interest.

The use of a plain and understandable language is essential in face to face services. The obvious advantage of clear communications is that people will easily understand the characteristic of the proposed service.

Therefore we simulated and intervention on the variable "Understandable language". We consider the optimal situation where language understandability is always evaluated 5. Such an improvement action on "Understandable language" has a significant effect on both "Professionalism" and the response variables; see Figure 8. This is mainly due to the strong association between "Understandable language" and "Product knowledge" that in turn is extremely relevant for the target variables, as shown in Figure 7. Notice that the node "Product knowledge" is on the path connecting "Understandable language" to "Professionalism" and to the target variables. Hence, "Understandable language" and "Professionalism" are independent given "Product knowledge". From a managerial point of view, if an improvement action is performed to increase employee product knowledge evaluation to the top score, actions on language understandability become ineffective.

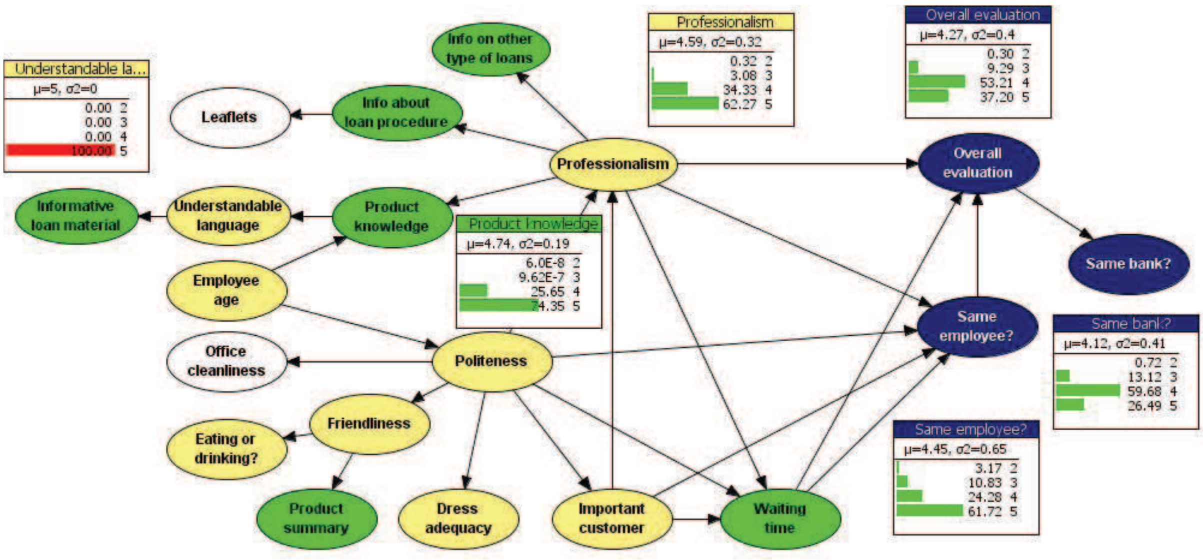

Figure 8: Effect of an improvement action on “Understandable language” and on the global evaluation of the service

The improvement scenarios described above are summarised in Table 1

Scenario E: Inverse Inference

The managers of the bank may be interested in getting some specific results for the target variables. In this case it is necessary to understand what are the quality aspects on which management has to focus in order to achieve the planned results. The network can be used also to answer this kind of question. Differently from the previous cases, evidence is inserted in the target variables and the updated marginal probability tables associated to the nodes directly or indirectly linked to the target ones are read and compared with the starting ones. In this way it is easy to identify the variables on which it is necessary to concentrate the managerial effort.

As an exemplificative scenario, consider the case in which the target is to avoid low evaluations (values 2 and 3) for variable “Overall evaluation”. After inserting this information in the network, we obtain the results shown in Figure 9.

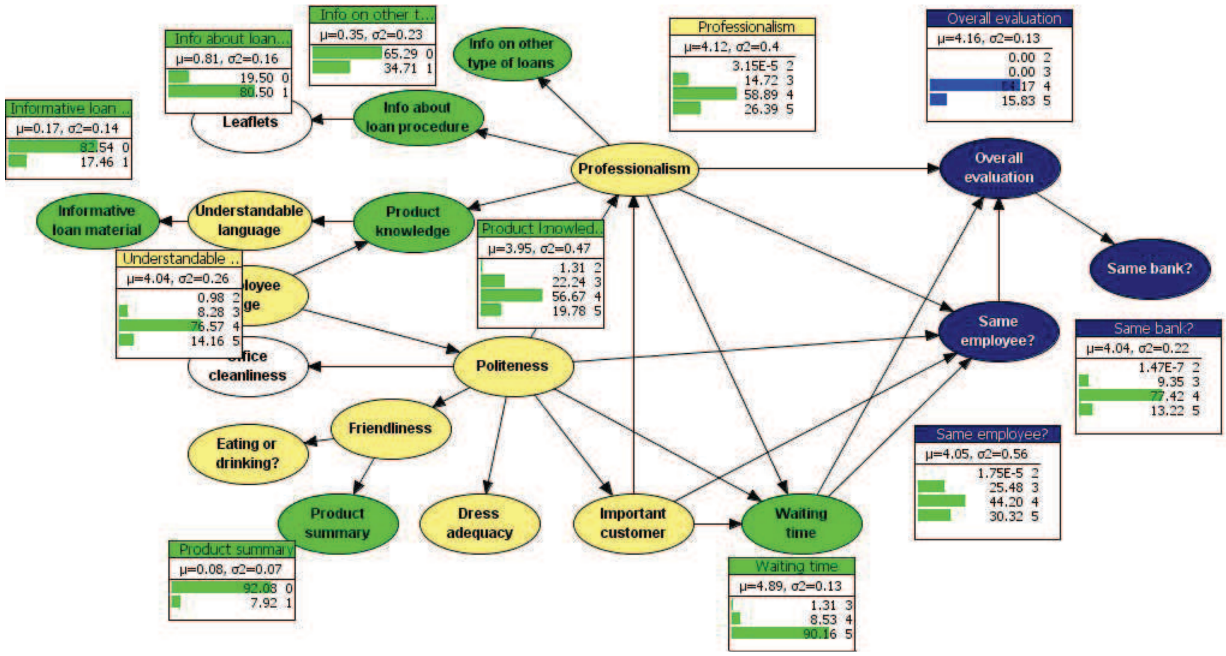

Figure 9: Network representing the optimal situation where “Overall evaluation” gets only scores 4 and 5

To get these results one should intervene, principally on variable “Professionalism” (with scores 4 and 5 rising from 60.02% to 85.28%), “Info about loan procedure” (employees should provide information on the loan procedure in 80.50% of the cases instead of 65.71%) and “Understandable Language” (with scores 4 and 5 rising from 80.55% to 90.73%).

Notice that this approach can be very useful also from an exploratory point of view. In fact, having in mind a managerial target, the network can be used to find out proposal scenarios whose specific effectiveness can then be evaluated as done in scenarios A–D. Here for example, following scenario D, it is enough to raise the highest scores of “Understandable language” and “Info about loan procedure” to achieve the goal of avoiding evaluations 2 and 3 for variable “Overall evaluation”.

4 Concluding Remarks

Service performance and its impact on the customer experience are key factors for bank management. Customers are free to choose between competitive alternatives, therefore companies should pay attention not only to the quality of service provided but also to its effectiveness. One method for service evaluation, that has increased in popularity in recent years, is the use of mystery shoppers. Mystery shoppers are "fake customers" used to survey and monitor the quality of the service and to identify areas requiring enhancement. After each visit they complete a report prepared in advance on their service experience, obtaining in this way a clear picture of strength and weakness of the service provided.

In this paper we have proposed BNs as a novel methodology for the analysis of mystery shopping data. The use of a BN allows to combine subject-matter knowledge and data derived information. BNs provide a structure that can be used for measuring and explaining overall customer satisfaction, and statistical methods to calculate the impact of different components on the overall satisfaction. Computationally efficient algorithms for evidence propagation in BNs are available; hence various possible improvement scenarios can be easily simulated and evaluated.

We have presented the results of an application of BN to mystery shopping data set regarding customer satisfaction of clients of Greek banks. In the application we have identified the key factors that influence global satisfaction of the clients, suggesting potential improvement areas for service production processes. The results of this analysis have showed that BNs are an efficient tool for service improvement analysis, considering customer perceptions. Using the information enclosed in BNs and the know-how concerning the bank organization, the managers can take decisions supported by a scientific and objective tool. To sum up, BNs not only describe the actual situation but allows to simulate, in real time, the effects of any improvement strategies.

Acknowledgements

The work of the first author was partially supported by MIUR, Italy, PRIN 2007XECZ7L, and the University of Pavia. The work of the second author was partially supported by MIUR, Italy, PRIN 2007BJK2PT, and the University Roma Tre. The authors are grateful to Mrs. Anna Sergianiti for providing us with the data and conducting the mystery shopping visits.

References

[1] Cheng, J.,Greiner, R., Kelly, J., Bell, D. and Liu, W. (2002), Learning Bayesian networks from data: An information-theory based approach, *Artificial intelligence*, **137**, 43-90.

[2] Chow, C. K. and Liu, C. N. (1968), Approximating discrete probability distributions with dependence trees, *IEEE Transactions on Information Theory*, **14** (3), 462-467 .

[3] Cooper, G.F. and Herskovits, E.A. (1992), A Bayesian method for the induction of probabilistic networks from data, *Machine Learning*, **9** , 309-347.

[4] Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. (1999), *Probabilistic Networks and Expert Systems*, Springer, New York.

[5] Grigoroudis, E., Politis, Y. and Siskos, Y. (2002), Satisfaction benchmarking and customer classification: An application to the branches of a banking organization, *International transactions in operational research*, **9**, 599-618.

[6] Kagara, E. and Voyiatzis, C. (2010), The customer service in the banking sector as a means of creating a competitive advantage: the case of Attica Bank, In Proceeding of *Management of International Business and Economics Systems 2010*, 644 - 657.

[7] Jensen, F. V, (1996), *An introduction to Bayesian Networks*, UCL Press, London.

[8] Lauritzen, S.L. (1995), The EM algorithm for graphical association models with missing data, *Computational statistics and Data Analysis*, **19**, 191-201.

[9] Lauritzen, S.L. (1996), *Graphical Models*, Oxford University Press, Oxford.

[10] Mihelis, G.,Grigoroudis, E., Siskos, Y., Politis, Y. and Malandrakis, Y. (2001), Customer satisfaction measurement in the private bank sector, *European Journal of Operational Research*, **130**, 347-360.

[11] Mylonakis, J. (2009), Customer Relationship Management Functions: A Survey of Greek Bank Customer Satisfaction Perceptions, *The IUP Journal of Bank Management*, 7-31.

[12] Neapolitan, R. E. (2004), *Learning Bayesian Networks*, Pearson Prentice Hall, NewHaven, Connecticut.

[13] Renzi, M. F., Vicard, P., Guglielmetti, R., and Musella, F. (2009), Probabilistic expert systems for managing information to improve services, *The TQM Journal*, 21, 429-42.

[14] Roberts, B. and Campbell, R.C. (2007), Being new-customer friendly: determinants of service perceptions in retail banking, *International Journal of Bank Marketing*, **25**, 56-67

[15] Saha, K. (2009), Mystery Shopping; A Marketing Research Tool to Measure Customer Satisfaction, *Journal of Business and Economic Issues*, **1**, 86-91.

[16] Salini, S. and Kenett, R. S. (2007), Bayesian networks of customer satisfaction survey data, *Journal of Applied Statistics*, **36**, 1177-1189

[17] Schrader, U. (2006), Ignorant Advice-Customer Advisory, Service for Ethical Investment Funds, *Business Strategy and the Environment* **15**, 200-214.

[18] Sergianniti, A. (2003), *A Mystery Shopping Survey in Greek Banks* , M.Sc. Thesis (in Greek), Department of Business Administration, University of the Aegean.

[19] Sherman, H. D. and Zhu, J. (2006), Benchmarking with quality-adjusted DEA (Q-DEA) to seek lower-cost high-quality: Evidence from a U.S. bank application, *Annals of Operational Research* , **145**, 301-319.

[20] Sherman, H.D. and Zhu, J. (2006), Benchmarking to Seek Lower-Cost High-Quality Service with Data Envelopment Analysis: Evidence from a U.S. bank application, *Annals of Operational Research*, **145**, 301-319.

[21] Steck, H. (2001), *Constraint-Based Structural Learning in Bayesian Networks using Finite Data*, PhD thesis, Institut für Informatik der Technischen Universität München.

[22] Wilson, A. (2001), Mystery shopping: Using deception to measure service performance, *Psychology Marketing*, **18**, 721-734

[23] Zikmund, W.G., Babin, B.J., Carr, J.C., Griffin, M. (2009), *Business Research Methods (with Qualtrics Card)*, Cengage Learning.

Appendix

Here we provide a short description of the variables considered in our network. For more details see Sergianniti (2003).

Black coloured nodes: Evaluation of the service

- *"Overall evaluation"*: Ordinal variable that provides a global evaluation of the service
- *"Same employee?"*: Binary variable indicating if the shopper would like to be served by the same employee again
- *"Same bank?"*: Binary variable indicating if the customer would like to be served by the same bank again

Dark grey coloured nodes: Service organisation aspects

- *"Waiting time"*: Ordinal variable indicating if the waiting time was considered adequate by the shopper
- *"Info on other type of loans"*: Binary variable indicating if the loan consultant informs the customer about alternative types of loan
- *"Info about loan procedure"*: Binary variable indicating if the loan consultant analytically informs the customer about the procedure and the subsequent action in case of a loan
- *"Product knowledge"*: Ordinal variable indicating if the loan consultant has a good knowledge of the products/services that he is offering
- *"Product summary"*: Binary variable indicating if the loan consultant recapitulates the products/services that were presented

Light grey coloured nodes: employee items

- *"Office Cleanliness"*: Ordinal variable indicating if the office of the loan consultant is clean and neat

- *“Dress adequacy”*: Ordinal variable indicating if the loan consultant has the appearance that is required by his work
- *“Employee age”* : Age of the loan consultant
- *“Politeness”*: Ordinal variable indicating if the loan consultant was polite or not
- *“Friendliness”*: Ordinal variable indicating if the loan consultant was friendly
- *“Eating or drinking”*: Binary variable indicating if any employee was eating or drinking during office hours
- *“Important customer”*: Binary variable indicating if the loan consultant encounter the client as an important customer
- *“Professionalism”*: Ordinal variable indicating if the loan consultant shows professionalism with his whole appearance
- *“Understandable language”*: Ordinal variable indicating if the loan consultant gives information in plain/understandable language

White coloured nodes: bank items

- *“Leaflets”*: Binary variable indicating if there are informative leaflets at the disposal of the clients of the bank
- *“Informative loan material”*: Binary variable indicating if the loan consultant gives relevant printed material to the client

Table 1: Comparison of Improvement Scenarios A–D

				TARGET VARIABLES					
				SAME EMPLOYEE		SAME BANK		OVERALL EVALUATION	
				Score	Percentage	Score	Percentage	Score	Percentage
ACTUAL SITUATION				5	18.19 %	4	52.11 %	5	9.50 %
						5	7.93 %		
SCENARIO	IMPROVEMENT ACTIONS								
		State	Percentage	Score	Percentage	Score	Percentage	Score	Percentage
A	*Product summary*	1	100%	5	62.71 %	4	60.78 %	5	33.60 %
	Informative loan material	1	100%			5	24.17 %		
		Score	Percentage	Score	Percentage	Score	Percentage	Score	Percentage
B	*Professionalism*	5	100 %	5	94.92 %	5	41.25 %	5	59.97 %
C	*Product knowledge*	5	100 %	5	76.88 %	4	57.93 %	5	47.26 %
						5	33.04 %		
D	*Understandable Language*	5	100 %	5	61.72 %	4	59.68 %	5	37.20 %
						5	26.49 %		