Michela Lacangellera and Paolo Mariani[1]

Main Features of Italian banking service: a cross-regional dynamic analysis

[1] University Bicocca -Milano, department of Business statistics;
ADDRESS: via Bicocca degli Arcimboldi, 8, 20126, Milano;
e-mail: michelalacangellera@yahoo.it; paolo.mariani@statistica.unimib.it.

Main Features of Italian banking service: a cross-regional dynamic analysis

Abstract: A strategy of exploratory multidimensional statistical analysis about key attributes of banking services is proposed. A static perspective, showing customers evaluation of banking offer as a whole, is followed by a dynamic view, analyzing regional paths across the considered waves, according to the importance and evaluation of the service attributes.

1. The reference context: an introduction

The Italian banking field, together with numerous institutional and economic changes in Europe, is evolving both in a prescriptive and in a business way.

Regarding the former aspect, the difficult role of banks and the increasing level of risk due, since the Seventieth on, to many considerable banking crisis, made the vigilance authorities define the minimum capital requirements for banks.

In fact, according to the agreements of Basel II, banks have to maintain a minimum level of capital in relation to risks that might really be run. The typologies of risk to consider for minimal capital requirements are: credit risk, market risk and operational risk.

On business side, banks are changing the features of their offer, so to be more and more considered as firms: managements problems and processes are rationalized, distribution channels and sale network are redesigned, and a new way to work is introduced.

In order to do that, companies try to grow according to logics and investments often not suitable to be taken simply by accounting techniques.

Furthermore, the pressing necessity to give an answer to the shareholders' expectations compels banks to be careful to the perception of their own choices and actions by the different subjects they relate to. Consequently, not only what's done and what's to be done are important, but even the perception that every shareholder has of it.

So, in order to let the interpretation of banks' actions be not arbitrary but in keeping with the former motivations and intentions, it's important to find out those means suitable to fully explain their own choices. For this reason, the most advanced companies decide to support the planning activity and the control of the financial-economic performance with some operating drivers tending to monitor continuously the
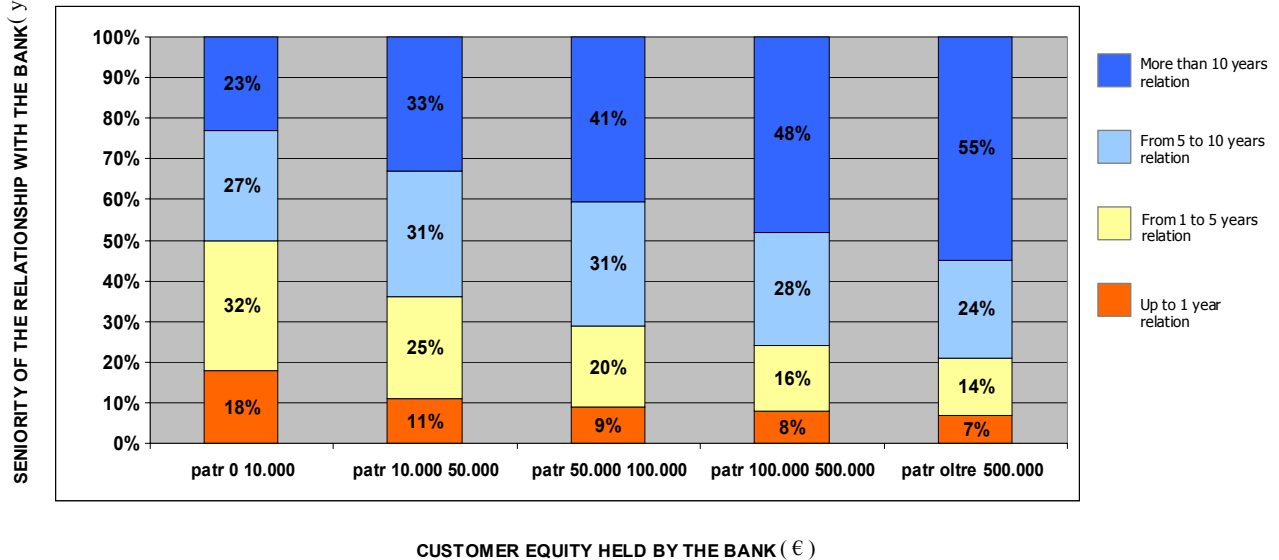
different shareholders' perception of the company's' actions, trying to point out how they come to judge the way to act.

In this context, the customer segmentation become more and more important for defining differentiated typologies of goods/services for every basin covered. We could say that, in Italy at least, banking offer is rather uniform (think about current accounts: the trade off between the lowest cost and the highest service quality makes certain typologies of current account offered by different banks more or less equivalent) so the differentiation seems to be the most powerful mean to maximize margins. The awareness of the most important attributes for a certain typology of customers, permits not only to obtain optimum conditions for the structure of banking costs (if an offer attribute seems not very important to a customer, it can stop to be offered in favour of other attributes), but also retain old clients and to attract new ones.

As illustrated by Graph 1, resulting from a research realized on the retail segment of the greatest Italian bank, it's just the old customer relation to create the highest profit. This is due both to the confidence component - the most important thing in a banking contest - and to the nature of the bank service - suitable to cross-selling activities for the old customers.

Graph 1:

Permanence of the relation and economic value of the customer

SENIORITY OF THE RELATIONSHIP WITH THE BANK (years)

	More than 10 years relation
	From 5 to 10 years relation
	From 1 to 5 years relation
	Up to 1 year relation

	patr 0 10.000	patr 10.000 50.000	patr 50.000 100.000	patr 100.000 500.000	patr oltre 500.000
More than 10 years	23%	33%	41%	48%	55%
From 5 to 10 years	27%	31%	31%	28%	24%
From 1 to 5 years	32%	25%	20%	16%	14%
Up to 1 year	18%	11%	9%	8%	7%

CUSTOMER EQUITY HELD BY THE BANK (€)

SOURCE: ABI DA BOCCONI MONITORING, *RETENTION E PERFORMANCE COMMERCIALI (2001)*

Therefore, in order to improve the offer, it's important to understand what to rely upon in accordance to what the customer requires as he has to choose between one operator and another.

For this purpose, one of the greatest Italian banking groups has made a research, in the national field, composed of three waves (one a year), with 27 thousands interviews a wave.

This research has been focused on the quality of the banking services across Italian regions, measuring the evaluation of main attributes of the services and their weight (importance) in the whole offer.

The submitted questionnaire is composed of 16 closed-questions (on a Likert scale 1 to 10) plus one final question on the global satisfaction of the service, and an initial part for personal data. Each question comprehends the so called "valuation of the service"

(type questions "B") – asking to valuate the item – and the so called *desiderata* about the service (type questions "A") – in order to express about the importance of the same item.

This questionnaire has been devised thinking about the main elements of customer satisfaction, that can be summarized in five clusters: tangible aspects, trustworthiness, respondence, capacity to reassure, empathy. Each cluster is well described by 3-4 questions A and B.

This analysis has considered not only the geographical segmentation – that will be used for the dynamic factor analysis- but even, for a descriptive phase here not reported, a segmentation by typology of customer now used in this group.

For this segmentation, every customer segment is defined by: age, average income, educational qualifications, work and typology of banking services used up.

After having outlined a framework of this analysis, the following step will be drawing a dynamic path, through the waves, of the main variables describing the banking service in the Italian regions, so to put in evidence some summarizing factors. It deals with summarizing the variability of a complex phenomenon by putting in evidence both similarities / dissimilarities among the "occasions" considered (the waves) and the main components of the average behaviour in the time interval chosen.

The techniques used to analyse the data "volumes", three way matrixes, are based on multivariate dynamic analysis.

Having only three waves like occasions, we preferred to use a compromised factorial plan rather than applying a dynamic factorial plan that, combining the principal component analysis and the multiple regression of the variables respect to the time (occasions), would need a longer number of observations to be reliable.

2. Multiway factor analysis: the methodology

Dynamic multivariate techniques of analysis make it possible to manage and analyse complex data structures, so to be able to study a given instance phenomenon in both a structural way – it means to fix basic relations among the objects (variables) we are interested in – and a dynamic way – in order to find out change and development of those objects in accordance to the occasions we are referring to.

We go to analyse three-way data matrixes individuals * variables * occasions, it means X_{ijk} type, with i=1,..,I individuals, j=1,..,J variables, k=1,..,K occasions. In our case, we consider Italian regions as individuals (so i=20), our variables are the 16 A and B variables plus the summarizing valuation on the satisfaction about the banking service as a whole (so j=33), the occasions are our waves (k=3).

This exploratory analysis is possible if:

1. Three-way data matrixes have same variables and same individuals for each occasion;

2. Three-way data matrixes have same individuals but different variables for one or more occasions (all variables depend on a specific *k* occasion);

3. Three-way data matrixes have same variables but different individuals for one or more occasions (all individuals depend on a specific *k* occasion).

The examined case is conforming with the first one, so we can operate on data matrix X_{ijk} directly.

According to k=3, it's built a common factorial space, the compromise space, in which the analysis elements are represented, according to their average configuration relative

to data volume as a whole. This space is obtained by means of a principal component analysis on the compromise matrix. This matrix is obtained by the weighted mean of the similarity/distance- matrixes among individuals[2].

The weighting coefficients are the eigenvector corresponding to the first eigenvalue of the similarity/distance matrix. Being based on the first eigenvalue only, the compromise matrix is robust, because not influenced from the small variations of the similarity matrixes. Therefore the compromise matrix is the synthesis of all matrixes, that are considered through the most representative average one. Through the factorial plan centred on this matrix, it's possible to examine the distance of the different matrixes from the compromise one (that is excluding the common part of variability), and their relation respect to the two principal axes.

On the axes obtained, we can draw both variables and individuals (our regions). This application ends with an analysis of the trajectories run by the units (individuals) respect to the compromise factorial plan.

3. The case

To act a global comparison among the occasions, correlation coefficients between the first, the second and the third wave have been studied.

The waves are highly correlated each other, that means a behavioural homogeneity as regards the considered instances, furthermore it justifies to continue the analysis and to search the compromise matrix. In fact, if there would be a high unhomogeneity among the occasions, the representation of the average behaviour in the considered years, will

---

[2] The similarity/distance- matrixes are matrixes which elements are the traces of the variance- covariance matrixes among the occasions.

risk to make the total variability of the phenomenon flat, biasing the points cloud. The high and significant correlation among waves is due to the fact that the one year time space between the one and the other is enough short to cause unhomogeneity in valuations and it's important for the reduction of information arising from the factorial analysis: axes will be enough solid to let represent in a well polarized way both individuals and variables.

At this point we proceed to the principal component analysis (PCA) to construct the compromise matrix, representing the average behaviours of the individuals in the three occasions considered.

In order to evaluate the quality of the factorial representation done, we report in Table 1 the three highest eigenvalues and the variability explained percentage on the total variability.

In addition to the initial solutions, the weight of the rotated factors are shown.

We operated an axes rotation by the Varimax method with Kaiser normalization[3], in order to increase the output interpretation. The Varimax method has been chosen because it was necessary to minimize the number of variables having high saturation per factor.

We can notice that the first three components explain more than the 80% of total variance of the system and the first factor by itself explains about the 43% of it.

---

[3] Varimax, which was developed by Kaiser (1958), is indubitably the most popular rotation method by far. For varimax a simple solution means that each factor has a small number of large loadings and a large number of zero (or small) loadings. This simplifies the interpretation because, after a varimax rotation, each original variable tends to be associated with one (or a small number) of factors, and each factor represents only a small number of variables. In addition, the factors can often be interpreted from the opposition of few variables with positive loadings to few variables with negative loadings.
Formally varimax searches for a rotation (i.e., a linear combination) of the original factors such that the variance of the loadings is maximized, which amounts to maximizing $V = \Sigma(q^2 j, l - q^{-2} j, l)^2$, with $q^2 j, l$ being the squared loading of the jth variable on the l factor, and $q^{-2} j, l$ being the mean of the squared loadings (for computational stability, each of the loadings matrix is generally scaled to length one prior to the minimization procedure). For Further information see Kim, Mueller 1978.

Therefore it seems reasonable, for the factorial representation, to consider variables and individuals respecting the first two factors.

Table 1

Initial Eigenvalues and rotated solutions

Component	Initial eigenvalues			Ratation of sums of squared loadings		
	Total	% of variance	% cumulative	Total	% of variance	% cumulative
1	14,121	42,791	42,791	14,097	42,719	42,719
2	10,550	31,971	74,762	9,733	29,495	72,214
3	2,574	7,800	82,562	3,339	10,120	82,334
4	1,044	3,164	85,726	1,120	3,393	85,726
. . .						

Extraction method: Principal Component Analysis

These results confirm what's emerged from the correlations between the waves: the high similarity among occasions comes to a high explanatory power of the PCA, making possible to summarize almost the total variability of the system.

Table2

KMO and Bartlett's test of sphericity

eiser Meyer Olkin Measure of Sampling Adequacy		,832
Bartlett's test of sphericity	Approx. Chi-Square.	3313,669
	df	528
	Sig.	,000

As a further confirmation of the quality of this variable reduction done, KMO and Bartlett's test of sphericity[4] are shown in Table 2.

---

[4]The Kaiser-Meyer-Olkin (1974) measure of sampling adequacy is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation

The first value is close to the unit (Kaiser, 1974, considered as satisfactory values higher than 0,7), pointing out that there are common factors of variability.

Furthermore Bartlett's test (1954) is very high with a very low significance level, rejecting the hypothesis of variables' independence.

Examining the variables contribution in axes construction, we can see from Graph 2, a clear polarization on both axes.
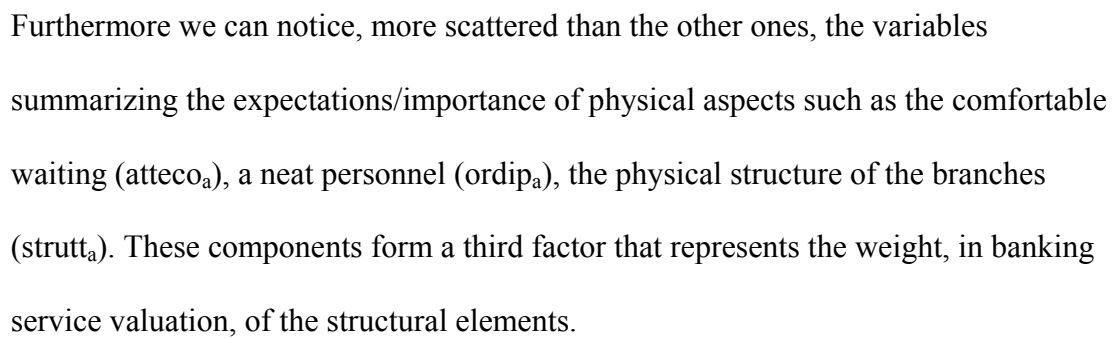
The first principal component is formed by variables type B for the positive extremity and by some variables type A on the negative side, vice versa for the second component. In particular the first one shows high positive correlations with variables that can express "valuations on the relational aspects and on the time component of the banking service", because these are variables type B (valuations) and deal with the comprehension of needs (compre$_b$), the rapidity of the service (serap$_b$), the competence and courtesy of the personnel (compet$_b$, corte$_b$), the punctuality (puntua$_b$) and the waiting time in queue (tempat$_b$).

The second axis indicates in a clearly the expectations of customers about the personnel competence (compet$_a$), convenient hours of the branches (oracom$_a$), punctuality, quickness of the service (puntua$_a$) and courtesy (corte$_a$). The importance of the temporal aspect outclasses visibly the other ones.

---

coefficients (refer to SPSS User's Guide). Large values for the KMO measure indicate that a factor analysis of the variables is a good idea.

Another indicator of the strength of the relationship among variables is Bartlett's test of sphericity (1954). It is used to test the null hypothesis that the variables in the population correlation matrix are uncorrelated. The observed significance level is .0000 (see Table 2). It is small enough to reject the hypothesis. It is concluded that the strength of the relationship among variables is strong. It is a good idea to proceed a factor analysis for the data.

Graph 2

A and B variables representation respect to the first and second factorial axes

Furthermore we can notice, more scattered than the other ones, the variables summarizing the expectations/importance of physical aspects such as the comfortable waiting (atteco$_a$), a neat personnel (ordip$_a$), the physical structure of the branches (strutt$_a$). These components form a third factor that represents the weight, in banking service valuation, of the structural elements.

Furthermore, the variable Valtot, expressing the global valuation of the whole service is highly and positively correlated with the variables type B above mentioned and lowly correlated with variables type A. So we could assess that the global valuation of the service is essentially referred to the valuation of relational and timing aspects, rather than to structural aspects like the structures of the offices (strutf), the waiting time in a comfortable location (atteco), etc.

According to the axes description, the first quadrant is characterized by both a high score attributed to the relational aspects of the service offered and a high score to importance of them. The regional branches (units) that will be positioned in this quadrant are the best scored in terms of valuation of relational and timing aspects and of the global satisfaction (component 1). Customers of these branches attribute much importance to those aspects of the service they evaluate greatly offered by the bank. We could argue that these branches are working well, improving the "right" aspect of the service, the most important for the customer.

The second quadrant represents a "dissatisfaction" about those aspects that are judged as fundamental by the customer; we could call this area the "to work a lot zone".

In the fourth quadrant there will be those aspects customers judge not important but offered by the bank at a high level, that are the relational and timing aspects. Although customers recognize that the bank takes a lot of care of these aspects, they weight more other thinks in their "utility curve".

Thanks to compromise space, we can analyse the evolution, across the three waves, of the region's position respect to the first two axes, that is respect to relational aspects of the banking offer and to the expectations on the temporal aspect.
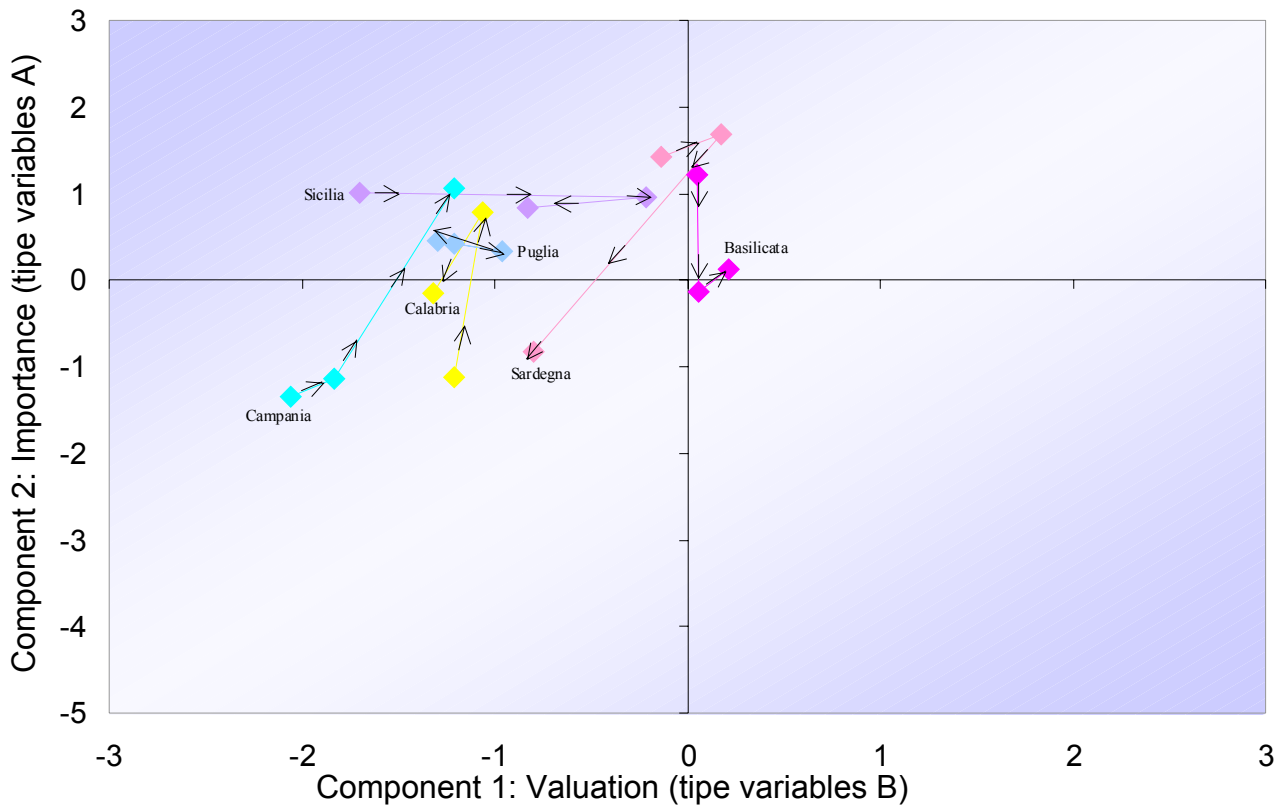
Graph 4:

Northern regions
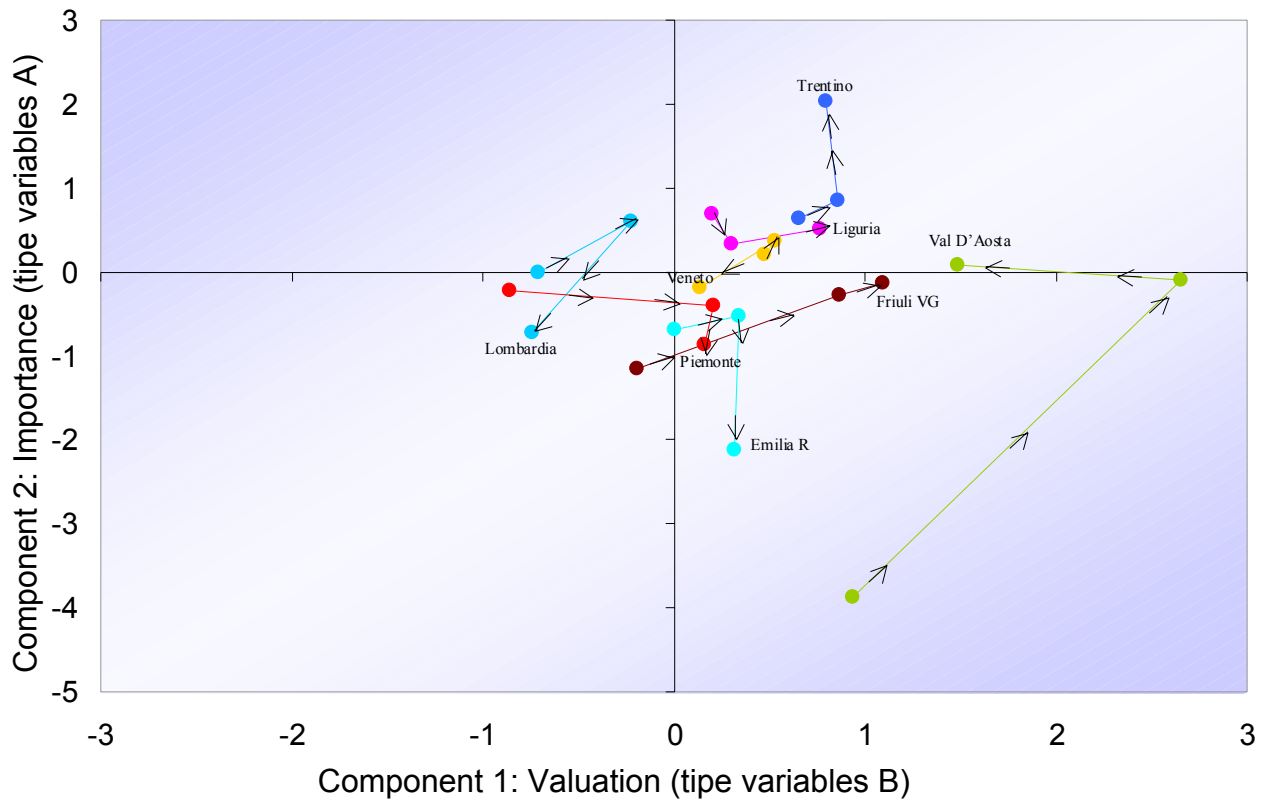
Graph 5:

Central regions
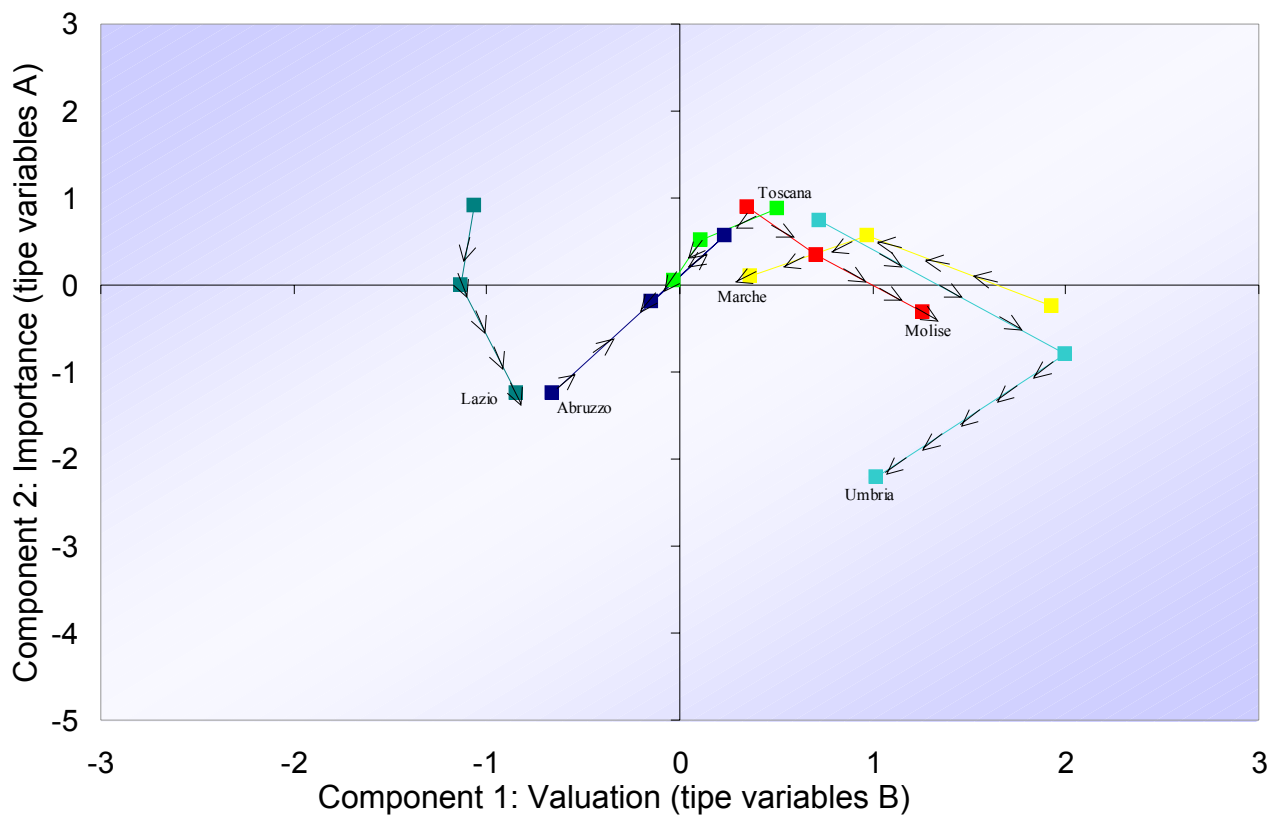
Graph 6:

Southern regions

For enhancing the intelligibility of the graphical representation we divided the 20 regions in three graphs: northern, central and southern regions, as shown in Graph 4, 5, 6.

Watching the three graphs we can say that, on the side of evaluation (first axis), the southern regions are concentrated in the second and third quadrants, the area that we called the "to work a lot zone". The movements across the wave don't improve significantly the initial position of each region.

Analyzing the movements of all the regions across the second axis, we can notice the regions that have incremented, both in the second and third wave, their requirement of shorter waiting time, more punctuality, more convenient hours of the branches, are

Trentino, Valle d'Aosta, Campania only. This occur, for another group of regions (Abruzzo, Calabria, Sardegna, Lombardia), only in the second wave. The remaining regions, who heavily (Umbria, Molise, Lazio) who in a light way (Toscana, Piemonte,…), maintained a trend of continuative decrease in the importance of punctuality, rapidity, but also personnel competence seems to be less crucial for customers of these regions.

For the interpretation of these results, we can add qualitative information. In fact the banking group reacted to the results of the first wave trying to improve satisfaction on those aspects judged important by customers (variables type A). For example the company instructed heavily the personnel and increased the number of employees. This made possible a reduction of waiting time jointly to an improvement of the problem solving capacity. In this way it is possible that relational and temporal aspects, from leading elements of satisfaction, became only constituent factors of it. For the first three regions mentioned, above all for Trentino and Campania, the increase of importance of such attributes can be explained by a significant increase of the affluent segment among customers.

Table 3

Weights of the key factors on the global valuation by customer segment

Key variables	Customer Segmentation %				
	Affluent	Traditional	Evolute	Retired	Univ. Sudent
Relationship quality	2%	26%	23%	21%	15%
Time	42%	29%	34%	26%	15%
Comfort	21%	26%	19%	28%	33%
Offices' structure	35%	18%	24%	25%	37%

From the descriptive analysis partly shown in Table 3, we can notice that this segment, probably for its life and work stile, attributes absolute importance to the temporal aspect in the whole offer and requires a high level of professionalism of the personnel to avoid waste of time.

In fact, the *traditional* segment comes to attribute a lot of importance to the relational aspects – like understanding requirements, confidence and staff's kindness- whereas the affluent segment gives top priority to the time aspects – like waiting in queue, punctuality, etc. Furthermore the so called *retired* segment attributes importance to the accessibility and comfort of the office.

Undoubtedly, these descriptive peculiar features recall the different life-stile of each segment.

Analyzing region's movements across the first axis, we focus on the real evaluation of the banking service (analysis of the same variables as before but type B).

We can notice that the major part of the regions in the three groups, from the first to the third wave, improve the valuation on waiting time and on personnel and also on the service as a whole (above all Campania, Piemonte, Molise, Liguria).

In particular these improvements are mainly concentrated between the first and second wave. We see a worsening in the third wave (Sicilia, Valle d'Aosta, Sardegna, Umbria e Lombardia) even if the whole effect, from the initial position, is positive. We can explain this trend, like for the second axis, as the reactions of the bank group to the results of the first wave. The improvements actuated by the bank enhanced the evaluation of the main attributes of the service.

At this point, to capture the real evolution of the service attributes according to customers, we can concentrate on the region trajectories only across the first axis. In

fact, the principal aim of the bank group is to maximize the evaluation of the attributes

by customers in a continuous trend across the waves. So we recognize two aspects to be

monitored: 1) The total path (that has to be the biggest possible): the sum of the

distances from the first wave to the second (w2-w1) and from the second to the third

(w3-w2) without a direction; 2) the net path (the most important aspect): the distance

from the first to the third wave (w3-w1) with the sign. If the net path is inferior to the

total path it means that region couldn't maintained the positive direction across the

waves: an improvement from a wave to another is partly offset by a worsening in the

other wave. This means that there is a "waste of time", in this sense it is not so

important "how many miles a region travels" (the total path), but "how

virtuous/efficient is the travel" (the net path). We can conclude that the best regions are

those that have no gap between total and net path and a positive direction, as shown in

Table 4. The most efficient path is that of Friuli, because it is the longest (1,28), it is

positive and there is no "waste of time" during the three waves (zero gap).

Table 4:

The paths across the waves

	Total Path	Net Path	Gap
Abruzzo	1,28	0,51	0,77
Basilicata	0,16	0,16	**0,00**
Calabria	1,40	-0,89	0,51
Campania	0,84	0,84	**0,00**
EmiliaR	0,36	0,32	0,04
FriuliVG	1,28	1,28	**0,00**
Lazio	0,35	0,22	0,13
Liguria	0,57	0,57	**0,00**
Lombardia	1,01	-0,03	0,98
Marche	1,56	-1,56	**0,00**
Molise	0,91	0,91	**0,00**
Piemonte	1,11	1,02	0,09
Puglia	0,58	-0,08	0,50
Sardegna	1,28	-0,66	0,62
Sicilia	2,10	0,88	1,23
Toscana	0,54	-0,54	**0,00**
Trentino	0,27	0,14	0,13
Umbria	2,26	0,30	1,96
ValdAosta	2,89	0,55	2,34
Veneto	0,45	-0,34	0,11

As we can see also in Graph 4, Val d'Aosta, in the third wave, offsets the major part of the improvement achieved from the first to the second wave: it's the most inefficient path. Marche and Toscana are the only regions with a negative continuous (no gap) path.

4. Conclusions

This investigation, though conditioned by the very limited number of occasions, tries to offer points of discussion on the evolution of customers' expectations and evaluations of a set of attributes of banking retail services. We can interpret the regions' trajectories in accordance to the effective actions taken by the company to meet customers need and in accordance to the particular situation of some regions.

The survey operated by the bank, as we showed in the paper, has been useful to guide bank's actions toward a better service from the first to the third wave. For this reason the management is continuing the survey with new waves. So the best follow up of this study will be an application of the real dynamic factor analysis, when the number of occasions will make possible the application of the multiple regression.

Furthermore this study could be extended considering, beside to these aspects, the customer segmentation that banks operate today at highest level of sophistication. In fact, it would be interesting studying how the weight of a particular segment in the bank clients' panel, is determining in the trajectory of one region: it seems reasonable that the importance of an attribute is very different for an affluent and for a traditional customer. In this sense we can easily notice the importance of a dynamic analysis, for example for a planner of a banking group, not only in a performance monitoring phase, but also in

assessing targets to branches. In fact, this kind of analysis shows immediately the strength and weakness of every phase of the supply of the banking retail service.

Bibliography

AIAF(2003), "La valutazione delle imprese bancarie: Un approfondimento del metodo EVA®", 114 (July), Milano.

Bolasco, Sergio(1999), *Analisi multidimensionale dei dati*. Roma: Carocci Editore.

Castillo Elizondo, William and Gonzalez Varala, Jorge (1994), "Anàlysis de tablas mùltiples de datos", *Revista de matematica: teoria y aplicaciones*, 1, 47-55.

Escoufier, Yves and Robert Pierre, (1976), "A unifying tool for linear multivariate statistical methods: the RV coefficient", *Applied Statistics*, 25,257-265.

Fraire, Mary (1989), "Problemi e metodologie statistiche di misurazione dei fenomeni complessi tramite indicatori e indici sintetici", *Statistica*, 2 (April-June), 246-263.

----and---- (2000) "Analisi dei dati a tre vie delle risposte a domande aperte e indicatori empirici", *Sociologia e ricerca sociale*, 61, 87-101

Kim, Jae-on and Mueller, Charles (1978), *Factor analysis: statistical methods and practical issues*. Thousand Oaks (CA): Sage Publications.

Lavit, Charles. and Escoufier Yves and Sabatier R. and Traissac Pierre (1994), "The Act (Statis Method)", *Computational Statistics and data Analysis*, 18, 97-119.