A Non-parametric Analysis of the Italian Banking System's Efficiency

Destefanis, Sergio

CELPE and DISES – University of Salerno

19 February 2001

A Non-Parametric Analysis of the Italian Banking System's Efficiency

Sergio Destefanis*
Università di Salerno

1. - Introduction

Many of the numerous works on technical efficiency analysis using production frontier calculations[1] have of late devoted much attention to non-parametric methods. These methods require a very limited number of hypotheses as regards the production process, as a producer's technical efficiency is assessed on the basis of production sets constructed by applying linear programming techniques without assuming the existence of a functional relation between input and output.

A distinction is usually made between those non-parametric methods which can be directly traced to the fundamental contribution of Farrell [12] (commonly grouped under the name of *Data Envelopment Analysis* or *DEA*) and those based on the *Free Disposal Hull* (*FDH*) approach which was first proposed by Deprins, Simar and Tulkens ([8]). This paper applies an extension of this latter method, which takes account of the existence of input and output slacks when calculating producers' technical efficiency, to the analysis of a

* The author, Researcher in Economic Statistics, is very grateful to Antonio Pavone for having generously made available his software for the non-parametric calculation of efficiency measurements. Obviously, the author bears sole responsibility for any errors or omissions.

N.B. the numbers in square brackets refer to the Bibliography at the end of the paper.

[1] Probably the most recent and complete reference on this is that of FRIED H.O. *et* AL. ([14]).

cross-section of 728 Italian banks belonging to different institutional categories (limited companies, credit societies, rural banks and credit co-operatives - SpA, banche popolari, casse rurali and casse di credito cooperativo respectively). Section 2 considers the main characteristics of the *FDH* approach and the nature of the extension (named corrected *FDH*, or *FDH-C*) used here. Section 3 contains a brief overview of the theoretical approaches to the analysis of the production process in banks, dwelling in particular on the definition of outputs and inputs. Section 4 briefly presents the data used while Section 5 illustrates the main results obtained, emphasising the territorial differences in technical efficiency, as well as any significance for efficiency of firm's size and of an environmental risk index. Section 6 presents some concluding considerations.

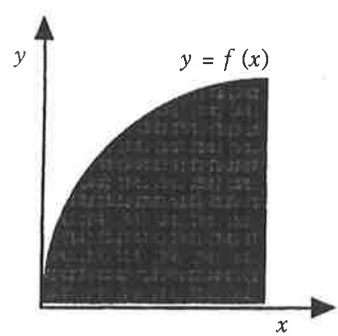
2. - Production Efficiency and *FDH* Approaches

The notion of efficiency of a production unit has been extensively dealt with in economic theory following the re-interpretation of the production function as a frontier of the set of production possibilities[2]. This set is defined by all the combinations of input and output which the producer can physically realise.

Consider for example Graph 1 which represents the usual production function $y = f(x)$:

GRAPH 1

PRODUCTION FUNCTION WITH A SINGLE INPUT AND A SINGLE OUTPUT



[2] See KOOPMANS T.C. [16], [17], DEBREU G. [7].

Each point of the production possibility set, characterised by an input-output pair, can be used to represent a production unit. The production function allows one to associate to each input quantity x the maximum possible amount of output y, so as to delimit the set of observations for the input-output pairs which characterise each production unit. It is in this sense that one speaks of the production function as a frontier and that the gap between a production unit and this frontier can be considered as a measure of its technical efficiency. More specifically, Koopmans ([16], [17]) proposes the following definition of technical efficiency; a unit engaged in the production of an output vector y by the use of an input vector x is technically efficient when the increase in an output necessarily implies either the reduction of another output or the increase of at least one input, and when the reduction of an input necessarily implies either the increase of another input or the reduction of at least one output. Hence a production unit is technically inefficient if it can obtain a given output vector after having reduced at least one input, or if it can increase at least one output using a given input vector.

In effect, we speak of: 1) input-oriented technical efficiency i.e., for a given level of output, the relation between the input corresponding to this output on the production function and the input actually used; 2) output-oriented technical efficiency i.e., for a given level of input, the relation between the output actually achieved and the output corresponding to this input on the production function.

Both parametric and non-parametric approaches are used in literature to derive a production frontier from data. In the former case, the frontiers of the production set are identified *a priori* with a production function, for example the Cobb-Douglas or the translog production function. If this function is estimated by maximum likelihood procedures, the difference between the production units observed and the values predicted by the production function can be broken down into two components: one purely random and one which represents the technical inefficiency of the units themselves[3].

[3] Indeed, if one uses the so-called stochastic parametric models, a given observation can be found above the production function as a result of purely random factors (which have nothing to do with the efficiency measurement). For an up-to-date overview of these methods, see FRIED H.O. *et* AL. [14]

Non-parametric approaches specify *a priori* not a given function, but rather some formal properties of the technology used (for example proportionality, convexity and free input and output disposal). Here again, the measure of efficiency shall be obtained with reference to the frontier of the production set. Starting from the work of Farrell [12] various linear programming methods have been developed to identify this frontier. Most of the procedures go under the name of *Data Envelopment Analysis* (*DEA*)[4]. Nonetheless, this paper will concern itself with a particular type of non-parametric approach known as the *Free Disposal Hull* (*FDH*)[5], based solely on the hypothesis of free input and output disposal. More specifically, for a given set of production units Y_0, the reference set $Y(Y_0)$ is characterised in terms of an observation i by the following postulate:

$$(x^i, y^i) \text{ observed, } (x^i + a, y^i - b) \in Y(Y_0) \quad a, b \geq 0$$

where a and b are respectively input and output disposal vectors. In other words, as a result of the possibility to dispose of input and output free of charge, the reference set includes all the production units which use larger or equal inputs and which produce lesser or equal outputs than i. The reference set can be indifferently a production set, a set of input requirements (for given outputs) or a set of outputs possibilities (for given input).

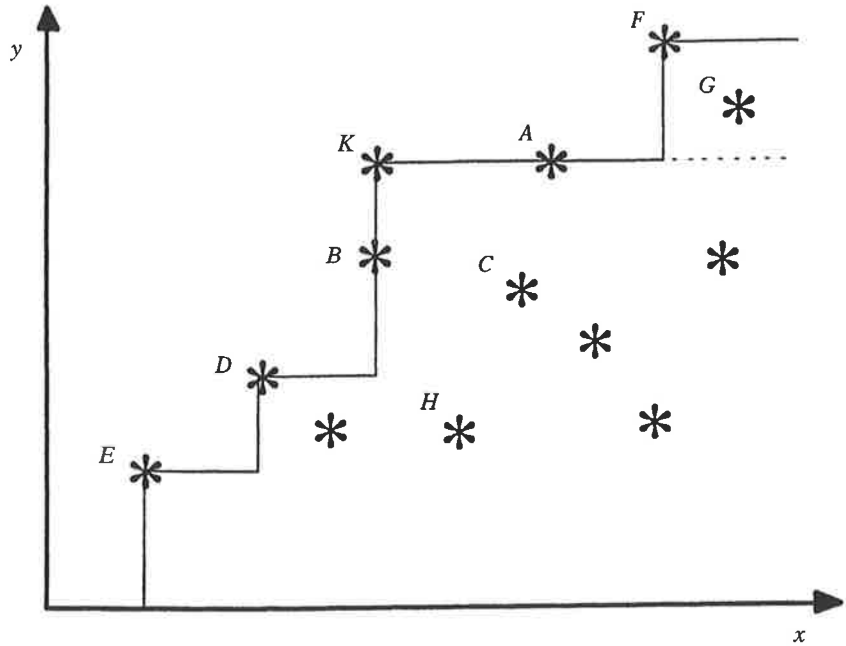
Take for example Graph 2 which considers a production set characterised by an input (x) and an output (y). Each input-output pair represents a production unit. Starting from the observation K, one defines every observation to the right and/or below it (e.g., with a greater input and the same output, as in A; or with lesser output and the same input, as in B; or yet again with a greater input and lesser output, as in C) as dominated by K. As regards G, it is dominated not by K, as it uses more input but al-

[4] See Charnes A. - Cooper W.W. and Rhodes E. [6] and subsequently Banker R.D. *et* Al. [3].

[5] A useful introduction to this method of analysis can be found in Deprins D. - Simar L. and Tulkens H. [8] and also in Tulkens H. [23], [24].

GRAPH 2

EFFICIENT AND INEFFICIENT PRODUCTION UNITS IN THE *FDH* APPROACH



so produces more output, but by *F*. On the other hand, *D* and *E* are not dominated by *K* and *F* as they produce less output but use even less input. Similarly, *F*, is not dominated by any other observation, as it uses more input but also produces more output. In fact, *K, D, E* and *F* are not dominated by any production unit belonging to the reference set.

The *FDH* approach makes this comparison for each observation, and the observations dominated by other production units are considered inefficient. On the other hand, those units not dominated by any other observation are considered efficient producers, and make up the frontier of the reference set. In order to measure the technical efficiency of the production units the Debreu-Farrell measure[6], either output- or input-oriented, is adopt-

[6] See DEBREU G. [7] and FARRELL M.J. [12].

ed. In the former case, the technical inefficiency (or, as it is commonly called, the efficiency score) equals the complement to unity of the maximum expansion of output in line with the use of a given input. A technically efficient producer (and hence on the reference frontier) cannot implement such an expansion of output, obtaining a efficiency score of one. As regards the input-oriented Debreu-Farrell measure, it is given by the complement to unity of the maximum reduction of input which allows one to maintain the production of a given output. In the case of a production unit, such as *H*, simultaneously dominated by two units on the reference frontier (*D* and *K*), to this unit is ascribed the efficiency score from the more dominant efficient observation (in the case in question *K* for output and *D* for input).

This type of analysis can be extended to the case of *n* dimensions in input and output. In this case, equiproportional expansions (of all outputs) or contractions (of all inputs) are considered in order to characterise production units situated on the production frontier and the complement to unity of the maximum equiproportional reduction of all the inputs which allows production of a given output vector to be maintained is taken as the measure of technical inefficiency. A technically efficient producer cannot implement such an equiproportional reduction of all outputs, obtaining an efficiency score of one. Similar considerations apply when the Debreu-Farrell measure (DF) is calculated as the complement to unity of the maximum equiproportional expansion of all outputs consistent with the use of a given input vector.

It is fundamental to note that in *FDH* an inefficient producer is necessarily dominated by at least one other specific producer (who really exists). This differentiates *FDH* from *DEA*, as the latter maintains that inefficient producers are dominated by virtual observations built as linear combinations of sets of efficient producers. The possibility, in *FDH*, to highlight some production units which actually exist, and to carry out direct comparisons between these and the units which they dominate, can be considered one of the greater merits of this approach.

Furthermore, the absence of any sort of assumption on the convexity of the production technology means that the frontiers obtained by *FDH* have greater probability of «being close to» the data than those obtained by *DEA*, when the reference set is characterised (at least locally) by the existence of non-convexity. Also as the frontier of the reference set consists of units which actually exist (rather than of a convex envelope), *FDH* is less sensitive to the presence in the reference set of anomalous (or wrongly measured) values than *DEA*. In fact, unlike with *DEA*, only the portion of the frontier corresponding to the anomalous value can be influenced by the latter's presence in *FDH*.

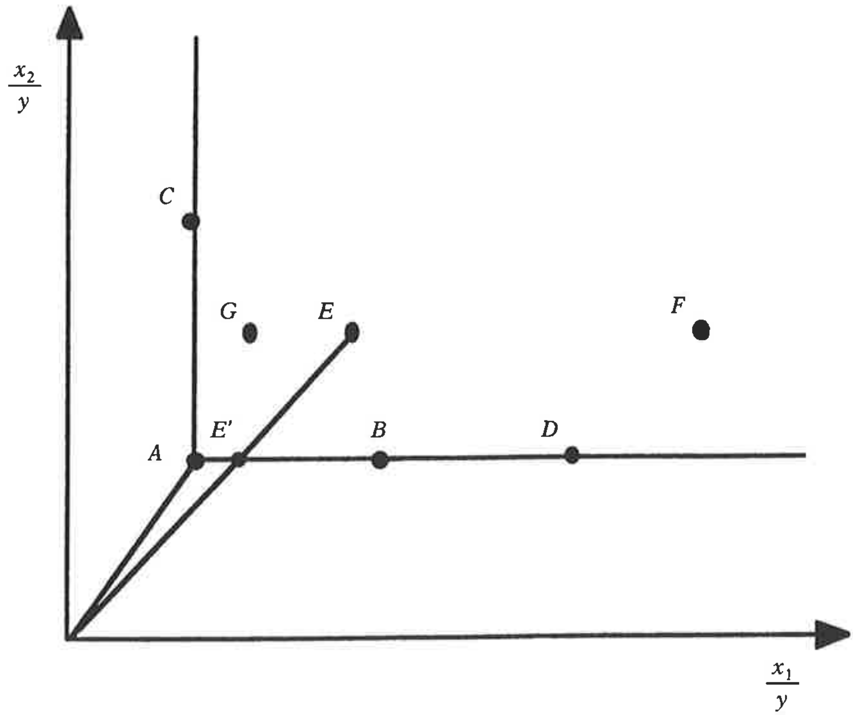
It is nonetheless important to emphasise how the definition of the reference frontier in *FDH* implies that a production unit can belong to it without dominating any other observation. Such a production unit (for example, *E*, in Graph 2) would be deemed efficient only by virtue of its location in an area of the reference set in which there are no other observations with which it can be compared. Following Tulkens's suggestion, such observations are defined efficient by default.

Furthermore, in *FDH* the efficient subsets coincide with the dominant observations. This drastically reduces the coincidence between efficient subsets and isoquants, and thus increases, for the same reason, the probability of discrepancies between Koopmans's definition and the DF measure of technical efficiency. In effect, a DF equal to one is a necessary but not sufficient condition for technical efficiency as defined by Koopmans. Intuitively, it is clear that the stepped profile of the *FDH* frontier is particularly suited to favour the existence of slacks, as is shown in the case with two inputs and one output depicted in Graph 3.

Graph 3 represents the set of input requirements for a unit with a technology comprising two inputs and one output. It can be easily ascertained that unit *A* is efficient even on the basis of Koopmans's definition. On the other hand, the input vectors of units *B*, *C* and *D* cannot be radially contracted allowing them to continue to produce the same output vector and nonetheless these units do not meet the criteria of technical efficiency definition provided by Koopmans. In fact, the quantity of input x_2 used by *C*

GRAPH 3

MEASUREMENT OF EFFICIENCY IN THE PRESENCE OF SLACKS



can be reduced without engendering any contraction in output. In other words, the input vector of unit *C* is characterised by the presence of slack in input x_2. The particular form postulated for the isoquant (which depends from its calculation using *FDH*) implies that the latter is equivalent only in part to a production subset which is efficient à la Koopmans. Given that DF requires only that an efficient vector lies on an isoquant, it follows that it cannot be a sufficient condition for Koopmans's criterion to be satisfied. Similar considerations apply with regard to the input x_1 used by *B* and *D*, as well as more generally, to the relations between the output-oriented DF measure and the definition of the same concept by Koopmans.

This example shows how it is easy in *FDH* to attribute the same DF result to production units with considerably different situations in respect of Koopmans's definition. It would therefore appear worthwhile to propose a measure of technical efficiency which takes account of the influence of slacks when assessing the efficiency of production units. The intuition which lies at the basis of the extension of *FDH* used here[7] is that slacks can be assessed in terms of radial contractions of input (or radial expansions of output) calculating a measure for the deviation of the input (or output) vector of an inefficient producer vis-à-vis the input (or output) vector of the respective dominant observation.

The analytical instrument used to measure this is the cosine of the angle formed by the input (or output) vector of the respective dominant observation and a virtual input (or output) vector obtained by radially contracting inputs (or radially expanding outputs) of a given observation until they reach the frontier of the reference set. Let us look at Graph 3 again. If we take observations *A* and *E* into consideration, the input vector of the latter will be contracted first to reach *E'*, and subsequently the cosine of the angle formed by *A* and *E'* will be considered as a measure of slack. The same procedure allows us also to evaluate the slack which characterises an observation on the reference set (such as for example *C*). In this case the measure of slack will be simply the cosine of the angle formed by *A* and *C*.

The measure of technical efficiency corrected for the presence of slacks, denoted DF-*C*, will be given for each production unit by the product of the radial DF measure for the cosine defined above, denoted *cos* (ω). In order to understand the reasons for the use of the cosine function, consider the formula of the coefficient of the correlation ρ between the two vectors x^A and x^B:

$$\rho = \frac{x^{A'} \; x^{B}}{[[x^{A}]] \; [[x^{B}]]}$$

[7] A more detailed presentation of this proposal is to be found in DESTEFANIS S. - PAVONE A. ([9]).

where $[[x^i]]$ denotes the Euclidean norm of the vector x^i. But the scalar product of the two vectors x^A and x^B can be defined as:

$$x^{A'} x^B = [[x^A]]\ [[x^B]]\, cos\ (\omega)$$

It follows that: $\rho = cos\ (\omega)$

The core of this proposal can therefore be summarised as follows. The more the composition of the virtual vector resembles that of the vector of the dominant observation, the closer the cosine to unity and the smaller the difference between the corrected and the original measure. In other words, the correction is a monotonically positive function of the amount of slack (whose size is given by the lack of correlation with the dominant vector). The greater the slack, the greater the correction applied to the radial efficiency measure for the respective producer.

It should be noted that in the field considered here, the function *cos* (ω) can never take zero or negative values. In fact, the function *cos* (ω) of any two observations takes the value zero only if at least two outputs or inputs alternately take zero value at each of the points of observation (the angle between the two points of observation therefore becomes a right angle). In this case however, it is impossible for one of the observations to dominate the other in the sense of the *FDH* approach. On the other hand, negative values of *cos* (ω) imply that at least one output or input for one of the points of observation takes a value of less than zero. But this is excluded by the production theory's standard hypothesis of non-negativity of quantities. Furthermore, a conseguence of the equivalence between cosine and the coefficient of correlation between the two vectors is that the correction proposed can be adopted whatever the number involved (greater than one)[8] of outputs and inputs. Also, it is easy to understand in the light of this very equivalence that the correction is not influenced by changes in the output and input measurement units.

---

[8] It is barely necessary to recall that in the presence of only one input and one output, the problem of the existence of slacks is resolved in an elementary manner by using non-radial efficiency measures.

Destefanis and Pavone [9] considered in detail the results of the application of *FDH* and *FDH-C* to the same data (local Italian hospitals for 1988, 1990 and 1991). These results show consistent and systematic differences in the efficiency scores assigned to the dominated observations which lead to quite significant alterations in the efficiency rankings obtained with the traditional *FDH* approach and suggest that *FDH-C* has an actual practical significance as regards ranking of the observations. This paper will therefore use this algorithm to measure the technical efficiency of a cross-section of Italian banks.

3. - The Definition of Input and Output in Banks

The seemingly unending amount of literature which has been dedicated to the production process in banks testifies to their central position in the economic system. We do not intend here to furnish an overview, however cursory, of these contributions which are numerous and concern themselves with various analytical fields. It would however appear appropriate to focus our attention on an aspect of this literature which perhaps more than any other concerns the analysis of technical efficiency, namely the definition of the inputs used and the outputs obtained within the ambit of this process. In this respect, it is useful to rely on a classification proposed in Berger and Humprey [4] who claim there are two main schools of thought in this respect: the asset approach and the value added approach.

The asset approach, which substantially coincides with the more traditional intermediation approach, emphasises the bank's role as a financial broker which transfers savings from units with a surplus to those with a deficit. Hence, the company's product consists of the assets of the bank, first and foremost the loans issued to the units with a deficit, while the deposits granted to the company by the units with a surplus, together with the other elements of the bank's liabilities, are the raw materials of the production process and should be included in the inputs along with labour and capital.

In more specifically taxonomical terms, this approach holds that the following should be considered as banks' outputs: the loan portfolio, which includes all the customer loans; the securities portfolio, which includes shares, bonds and shareholdings; the interbank market loans. As regards inputs, we can consider as raw materials customer deposits, Banca d'Italia funding, loans obtained on the interbank market. The other inputs are capital, consisting of premises and equipment, the so-called free capital, namely the equity funds net of fixed assets, and labour, represented by the number of hours worked.

The principal characteristic of this approach is naturally the inclusion of deposits and other liability elements among the inputs. Now, even if this would seem to capture some fundamental elements which characterise the bank's production process, there are nonetheless strong theoretical perplexities in this respect, particularly as regards the inclusion of customer deposits among the inputs. It would, in effect, appear difficult to exclude the latter from the outputs, when their management generates a significant part of many banks' operating profit. Moreover, it is difficult to see how a bank's customers would be prepared to pay a price (in the form of bank commission) on something which is not part of a bank's product.

In the value added approach, which draws to a considerable extent on the fundamental elements of the so-called production approach, these perplexities are taken to their extreme logical consequences and customer deposits are included among banks' outputs. According to this approach, banks provide services to customers on both asset and liability sides (loans, deposits, bank and credit cards, etc.) using basically labour and capital as inputs. More specifically, the value added approach holds that if an asset or liability item absorbs a significant amount of labour and capital resources then it is included among the outputs. Otherwise, it should be included among the inputs or among the non-essential outputs. Hence the outputs include current accounts, savings accounts, commercial credit, loans and other banking services such as foreign exchange intermediation, tax collection and securities broking and placement activities. Among the inputs, on the other hand, we include Banca d'Italia funding and large-sum certificates of deposit as their acquisition requires no significant use of labour and capital. Finally, the non-essential outputs are assets which do not entail significant use of labour and capital such as for example portfolio Treasury notes.

Table 1 summarises the input and output variables considered in some recent non-parametric analyses of banks' production efficiency.

Table 1 clearly shows how both the approaches illustrated above are represented among these studies. Drake and Weyman-Jones [10] and Elyasiani and Mehdian [11] espouse essentially the asset approach, while Aly *et* AL. [1], [2], Ferrier and Lovell [3], Fried *et* AL. [15], Tulkens [25] and Resti [21] can be substantially classified as embracing the value added approach[9]. In fact, it cannot be stated that there currently exists a consensus in literature on which of the two schools of thought is preferable *a priori*. Rather, as Resti noted ([21], p. 275), it is the aims of the empirical investigation and the concrete availability of data which determine the choice of the paradigm to be used.

As will become clearer below, in this case the characteristics of the database available advise the adoption of the value added approach. In any case, a couple of concluding considerations can be made which are valid for both approaches. Firstly, in both of them the banks' production process is characterised not only by a number of inputs, but also by a number of outputs which are heterogeneous to each other, thus making it difficult to obtain a single scalar output index. As a result, the analysis of production efficiency for these institutions is considerably improved by the application of non-parametric techniques such as *FDH-C* which make it comparatively simple to take into account multi-input, multi-output production technologies. Secondly, as is shown by the study by Fried *et* AL. [15], it is extremely important in principle to include some index of the quality and range of banking products offered by the various institutions as outputs. In the absence of such,

[9] It should be noted that some of these authors - FERRIER G.D. - LOVELL C.A.K. [13]; FRIED N.O. *et* AL. ([15]), TULKENS H. [25], use as output data the numbers of bank transactions (number of current accounts, loans, etc.) and not their respective asset value, thus being closer to the traditional production approach.

TABLE 1

NON-PARAMETRIC ANALYSIS OF EFFICIENCY:
PROCEDURES AND VARIABLES

Authors	Procedure used	Output variables	Input variables
ALY *et* AL. [1] [2]	*DEA*	- loans - deposits	- number of employees - fixed assets - free capital
FERRIER-LOVELL [13]	*DEA*	- no. demand deposits - no. time deposits - no. loans - no. instalment loans - no. commercial loans	- number of employees - fixed assets - expenditure for materials
DRAKE-WEYMAN-JONES [10]	*DEA*	- mortgage loans - other assets - excess liquidity	- no. of employees - capital at book values - retail deposits - wholesale deposits - no. of branches
ELYASIANI-MEHDIAN [11]	*DEA*	- no. of commercial loans - no. of mortgage loans - other loans - portfolio securities	- no. of employess - fixed assets - demand deposits - time deposits
TULKENS [25]	*FDH*	- no. of current account operations - no. of ATM operations - no. of foreign currency operations - no. of brokerage operations - no. of loans issued - no. of new accounts opened - no. of credit cards issued - other operations	- hours worked - no of branches - no. of ATMs
FRIED *et* AL. [15]	*FDH*	- no. of loans - loan range index - loan interest rates - no. of deposits - deposit range index - deposit interest rates	- no. of employees - other operating costs
RESTI [21]	*DEA*	- loans - funds raised from customers - net interbank loans	- no. of employess - no. of branches - net interbank funds - large-sum certificates of deposit

the efficiency gaps measured by empirical analysis could in fact derive from the differing characteristics of the output offered by the banks surveyed. Should it not be possible to have such indexes, it would therefore be advisable to avoid comparing banks belonging to different institutional categories and which therefore in all likelihood have a different composition of output.

4. - An Application for Banks. The Data

The data used in this paper refers to a sample of 728 Italian banks and is taken from the 1994 Rating data archive. The variables available for the afore-indicated sample are: total assets, total customer loans, total customer deposits, total securities funds raised (certificates of deposit and bonds), net interbank relations, equity, free capital, number of employees, number of branches, interest margin, labour costs, operating income, total bad loans. The banks taken into consideration were 505 rural banks (casse rurali) and credit co-operatives (casse di credito cooperativo) and 223 limited banks, credit societies and savings banks (SpA, banche popolari and casse di risparmio respectively) giving a total of 728 banks distributed throughout Italy. We eliminated from the original sample of 763 commercial banks 20 banks which stated they had zero employees, 2 banks which stated they had no branches, 4 banks which declared that funds raised from securities (certificates of deposit and bonds) amounted to zero and 9 banks with negative free capital, as the algorithm used for *FDH-C* can function only with positive variables. It is in any case reasonable to assume that these exclusions have not only not significantly diminished the informational content of the cross-section, but that they have also allowed the elimination of some potentially anomalous observations.

Some shortcomings can be noted if the variables available are compared with those listed in the previous section, above all as regards the asset approach. The Rating sample does not include data on Banca d'Italia funding. Furthermore, it is impossible to separate the portfolio securities from the total assets or to separate the total of securities funding raised into that raised from bonds and that from certificates of deposit (which are much more brokerable funds than bonds). These shortcomings in the data will also negatively influence the possibility of specifying the elements of a production set under the value added approach, but will be less serious if the securities portfolio management does not require a significant use of labour and capital and this asset item can therefore be considered as an unimportant output[10]. It should be noted in any case that the lack of data on the number of bank transactions (number of current accounts, loans, etc.) prevents the use of the value added approach in its form which is closest to the traditional production approach.

As is the case for most of the studies considered in Table 1, the Rating sample contains no data regarding the number of hours worked and as a result the labour input has to be measured by the number of employees. Furthermore, no measurements are available for banking services, such as foreign exchange intermediation, tax-collection, and securities broking and placement activities. As this absence also makes it impossible to construct appropriate indexes for the quality and range of banking products, it would appear advisable to compare only banks which belong to the same institutional category. In this case, considerations regarding the size of the sample counsel subdividing the sample into rural banks and credit co-operatives on the one hand, and limited banks, credit societies and savings banks on the other.

As Table 1 also shows, literature usually employs two types of variables to measure fixed capital: the book value of premises and equipment and the number of branches. We believe it appropriate to use the latter variable here as it is crucial for banks which aim at financial intermediation via a decentralised structure. On the other hand, it is impossible to separate the book value of premises and equipment from that of long-term investments in the Rating databank, and in any case the book value of fixed capital can be subject to errors of measurement deriving from the differing balance-

---

[10] For an analysis in favour of this hypothesis, at least for portfolio Treasury notes, see ONADO M. ([20], Chapter 10).

sheet criteria adopted (particularly relevant for comparisons between limited banks, credit societies and savings banks).

Finally, the Rating sample includes data only on net interbank flows, and not on the relative gross flows. Given that net interbank flows have a negative value in a hundred or so cases, to obtain a sample without negative values one would have to significantly reduce the cross-section by number. Moreover, in this case the conceptual requisites for believing that the observations to be excluded are potentially anomalous do not exist. In view of the lack of emphasis on this variable in literature – only Resti [21] includes net interbank flows in the analysis – it would appear opportune to maintain the cross-section of 728 banks in the empirical analysis, and disregard the net interbank flows when specifying the production sets.

In conclusion, it appears possible to state that the Rating database includes variables which enable it to be used for the empirical specification of an appreciably complete version of the value added approach. In effect, conditional on the assumption that the management of portfolio securities does not entail significant use of labour and capital, this database includes the more significant outputs for the value added approach (namely deposits and loans). As regards the possibility of adopting the asset approach, it should be noted that the reliability of the efficiency measurements based on this approach could be considerably reduced by the lack of a sufficient breakdown of the asset and liability items.

5. - An Application for Banks. The Empirical Analysis[11]

After some brief introductory comments, this section will illustrate the main results obtained by applying *FDH-C* to banks of the Rating sample. According to the value added approach, the outputs will include the total of customer loans and customer deposits. No account is taken of the rest of the assets, which, in ad-

[11] The efficiency measurements based on *FDH-C* were obtained by the software developed and kindly made available by Antonio Pavone. The Kruskal-Wallis and Tobit analyses were made using the Systat and Limdep software respectively.

dition to the capital, include portfolio securities that, as already noted, should not be essential outputs as understood by the value added approach. The inputs include not only the number of employees and branches but also the free capital. In effect, this variable, although it explicitly figures only once in the inputs presented in Table 1, is part of a company's funds and is traditionally considered a factor of some importance in respect of a bank's capacity to grant loans. On the other hand, it has already been used with good results in the parametric efficiency analysis of Locatelli and Prosperetti [18]. Finally, it should be borne in mind that it is not possible to include Banca d'Italia funding among the inputs, nor separate the funds raised from bonds and from certificates of deposit (which are much better brokerable funds than bonds). It would appear appropriate however to examine the possibility that the production set contains also the total funds raised by securities as an input. We will therefore have: the following production set according to the value added approach: 1) output: total loans and deposits; 2) input: number of employees, number of branches, free capital, total securities funds raised.

Table 2 shows some descriptive statistics for these variables which highlight the strong structural differences for the two big categories of institutions considered here.

Recall that the impossibility of building suitable indexes for the quality and range of banking products advised against comparing banks belonging to different institutional categories (on the one hand rural banks and credit co-operatives and on the other limited companies, credit societies and savings banks). These considerations are corroborated by the strong structural differences shown in Table 2. Hence, our efficiency analysis will reflect solely the outcome of the comparisons within the two main institutional categories, henceforth referred to as CRC and SPA respectively.

Table 3 shows the main results of the analysis which deals with input-oriented efficiency measuremet only. As can be seen, two different versions of the production set have been taken into consideration. In addition to employees and branches, account has also been taken respectively of free capital, and of this variable together with the total of securities funds raised. The criteria

TABLE 2

DESCRIPTIVE INPUT AND OUTPUT VARIABLE STATISTICS

SPA no. = 223		CRC no. = 505	
Loans			
Lower quartile:	255.345	Lower quartile:	21.646
Median:	689.522	Median:	46.170
Upper quartile:	1.825.502	Upper quartile:	85.809
Customer deposits			
Lower quartile:	333.874	Lower quartile:	37.107
Median:	867.094	Median:	69.844
Upper quartile:	2.036.705	Upper quartile:	129.063
No. of employees			
Lower quartile:	145	Lower quartile:	11
Median:	401	Median:	23
Upper quartile:	954	Upper quartile:	46
No. of branches			
Lower quartile:	13	Lower quartile:	2
Median:	31	Median:	3
Upper quartile:	68	Upper quartile:	5
Free capital			
Lower quartile:	35.728	Lower quartile:	4.932
Median:	88.879	Median:	9.913
Upper quartile:	215.643	Upper quartile:	18.746
Funds raised from securities and certificates of deposit			
Lower quartile:	117.074	Lower quartile:	11.662
Median:	291.125	Median:	22.943
Upper quartile:	768.662	Upper quartile:	47.745

adopted for comparing the different results are the average efficiency score of the dominated units and the number of dominant and efficient by default units obtained according to the type and number of inputs included in the production set. The higher the average efficiency score of the dominated units and the higher the number of dominant units, the greater the role played by the type and number of inputs included in the production set in explaining the production performance of the units considered, while a

TABLE 3

PRINCIPAL RESULTS OF THE *FDH-C* ANALYSIS

CRC (505 observations)			
Output: Loans, deposits			
Input: Employees, branches, free capital			
Number of dominant observations:		157	
Number of observations efficient by default:		80	
Average efficiency for the dominated observations:		0.888	
Sum of slacks for dominated observations:			
For output (*s*)		For input (*s*)	
Expansion in %		Reduction in %	
Loans	22	Employees	6
Deposits	15	Branches	23
		Free capital	18
Output: Loans, deposits			
Input: Employees, branches, free capital, securities funds raised			
Number of dominant observations:		158	
Number of observations efficient by default:		155	
Average efficiency for the dominated observations:		0.878	
Sum of slacks for dominated observations:			
For output (*s*)		For input (*s*)	
Expansion in %		Reduction in %	
Loans	17	Employees	7
Deposits	20	Branches	23
		Free capital	18
		Sec. funds raised	26
SPA (223 observations)			
Output: Loans, deposits			
Input: Employees, branches, free capital			
Number of dominant observations:		50	
Number of observations efficient by default:		92	
Average efficiency for the dominated observations:		0.877	
Sum of slacks for dominated observations:			
For output (*s*)		For input (*s*)	
Expansion in %		Reduction in %	
Loans	21	Employees	5
Deposits	16	Branches	19
		Free capital	37
Output: Loans, deposits			
Input: Employees, branches, free capital, securities funds raised			
Number of dominant observations:		31	
Number of observations efficient by default:		144	
Average efficiency for the dominated observations:		0.884	
Sum of slacks for dominated observations:			
For output (*s*)		For input (*s*)	
Expansion in %		Reduction in %	
Loans	18	Employees	5
Deposits	19	Branches	17
		Free capital	46
		Sec. funds raised	19

higher level of units which are efficient by default indicates that the inclusion of a given type and number of inputs only increases the non-comparability of the observations.

Consideration of these criteria leads to the result that for both the CRCs and the SPAs, the production sets which would appear to best account for the production performance of the units considered is that which takes employees, branches and free capital as inputs. In effect, the use of the securities funds raised together with free capital engenders no positive effects as the average efficiency score of the dominated units and the number of dominant units show only a slight increase or even decrease, while there is a significant increase in the number of units which are efficient by default. Furthermore, in both cases, the average efficiency scores are very similar. Finally, note that the results using the value added approach would appear to indicate that the slacks' input structure is heavily weighted towards the capital inputs. More specifically, slack would appear to be particularly pronounced in the CRCs as regards the number of branches. This result indicates that the proliferation of new branches after 1990 resulted in overbranching rather than overmanning for these banks (Resti [21], p. 295).

To analyse the characteristics of the efficiency scores in more detail, it would appear appropriate to focus on the production sets in which the inputs consist of employees, branches and free capital. To this end, a non-parametric analysis of the variance for the efficiency scores was first carried out, using the Kruskal-Wallis procedure (robust vis-à-vis the non-normality of the residuals) to examine the territorial distribution of the scores by macroregion and region. The existence of systematic patterns in the efficiency scores was then assessed using the Tobit analysis. A model with a limited dependent variable (such as Tobit) was used for the econometric analysis of the efficiency scores because, by definition, the efficiency score has a range of variation of between zero and one; ignoring these constraints on the variability of the efficiency scores could lead to distorted and inconsistent estimates. An attempt was also made here to link the distribution of the efficiency scores to two variables: the size of the company (measured by L(Att), the logarithm of the value of assets) and the riskiness of the environment (measured by the bad loans/total loans ratio, (SOFF%). Significance on the part of the former variable should indicate the existence of scale effects while the latter variable should capture the effect of the «quality» of the environment in which the bank operates.

One *a priori* consideration regarding the results is that the analysis of the territorial differences in efficiency is more informational in nature for the CRCs as they tend not to have branches outside their «home» region. In fact, the Kruskal-Wallis analysis, shown in Tables 4 and 5, indicates the existence of systematic elements in the territorial distribution of the efficiency scores solely for the CRCs. This indication in corroborated by the Tobit analysis of the relative patterns, which is shown in Tables 6 and 7. For the CRCs, the regions of mainland southern Italy are on average less efficient, while the SPAs would appear to have no significant territorial pattern. Note that to avoid the so-called multicollinearity trap, it was necessary to subsume a macroregion or region in the constant, which in the estimates shown here was always the median macroregion or region. Hence, the coefficients of the territorial dummies represent the gap of the coefficient of a given macroregion or region vis-à-vis that of the median macroregion or region.

There were significant differences between the CRCs and SPAs also as regards the role of the size of the bank and the riskiness of the environment. In the case of the SPAs, neither of the variables gave rise to any significant effects, while for the CRCs there is evidence of (albeit rather weak) scale effects and SOFF% has a negative coefficient and reduces the significance of the dummies (of a negative sign) for southern Italy. This type of evidence is compatible with the existence of a direct relation between problems of «quality» of the environment of southern Italy and the level of inefficiency of southern Italian CRCs[12]. Naturally, this point and the significance of the size variable both require further analytical investigation.

---

[12] There are *a priori* reasons for believing that the co-operative structure and the more informal procedures employed in lending decisions make the CRCs more sensitive to the environment in which they operate. See on this matter CANNARI L. - SIGNORINI L.F. [5].

TABLE 4

EFFICIENCY SCORES DF-C, FOR DOMINATED OBSERVATIONS
CRC (505 OBSERVATIONS)
ANALYSIS OF THE KRUSKAL-WALLIS VARIANCE*

North-West	N = 79	Emilia Romagna	N = 41
North-East	N = 257	Tuscany	N = 39
Centre	N = 80	Umbria	N = 4
South (mainland)	N = 69	Marche	N = 20
Sicily-Sardinia	N = 20	Lazio	N = 15
Piedmont	N = 12	Abruzzo	N = 13
Valle d'Aosta	N = 3	Molise	N = 3
Lombardy	N = 62	Campania	N = 22
Liguria	N = 2	Puglia	N = 15
Trentino-Alto Adige	N = 138	Basilicata	N = 5
Veneto	N = 56	Calabria	N = 11
Friuli-Venezia Giulia	N = 22	Sicily	N = 20

* Groups: North-West, North-East, Centre, South mainland, Islands.

K-W statistic = 9.29, *P*-value = 0.05 $\sim \chi^2$ (4)

Groups: Piedmont, Valle d'Aosta, Lombardy, Liguria, Trentino-Alto Adige, Veneto, Friuli Venezia Giulia, Emilia Romagna, Tuscany, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicily.

K-W statistic = 24.54, *P*-value = 0.14 $\sim \chi^2$ (18)

TABLE 5

EFFICIENCY SCORES DF-C, FOR DOMINATED OBSERVATIONS
SPA (223 OBSERVATIONS)
ANALYSIS OF THE KRUSKAL-WALLIS VARIANCE*

North-West	N = 57	Tuscany	N = 17
North-East	N = 59	Umbria	N = 8
Centre	N = 50	Marche	N = 7
South (mainland)	N = 40	Lazio	N = 15
Sicily-Sardinia	N = 17	Abruzzo	N = 8
Piedmont	N = 16	Molise	N = 1
Lombardy	N = 36	Campania	N = 11
Liguria	N = 5	Puglia	N = 14
Trentino-Alto Adige	N = 7	Basilicata	N = 3
Veneto	N = 15	Calabria	N = 6
Friuli-Venezia Giulia	N = 9	Sicily	N = 15
Emilia Romagna	N = 28	Sardinia	N = 2

* Groups: North-West, North-East, Centre, South mainland, Islands.

K-W statistic = 5.56, *P*-value = 0.25 $\sim \chi^2$ (4)

Groups: Piedmont, Valle d'Aosta, Lombardy, Liguria, Trentino-Alto Adige, Veneto, Friuli Venezia Giulia, Emilia Romagna, Tuscany, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicily.

K-W statistic = 21.98, *P*-value = 0.23 $\sim \chi^2$ (18)

TABLE 6

A TOBIT ANALYSIS OF THE EFFICIENCY PATTERNS
CRC (505 observations)

Dependent variable = efficiency score from model
Output: Loans, deposits
Input : Employees, branches, free capital

Variable	Coefficient	*t*-ratio	Variable	Coefficient	*t*-ratio
Log-likelihood = –79.02198			*Log-likelihood = –73.46187*		
Constant	1.0175	46.690	Constant	0.89962	7.289
North-West	0.59803E-02	0.196	North-East	–0.33342E-02	–0.133
North-East	0.75418E-02	0.308	Centre	0.87418E-02	0.285
South (mainland)	–0.64328E-01	–2.109	South (mainland)	–0.19898E-01	–0.582
Islands	0.29877E-01	–0.640	Islands	0.45041E-01	0.858
			L(Att)	0.12381E-01	1.211
			SOFF%	–0.71875E-02	–3.164
Log-likelihood = –68.85028			*Log-likelihood = –62.55388*		
Constant	1.0009	35.119	Constant	0.74043	5.586
PIE	–0.10975E-01	–0.183	PIE	–0.50021E-01	–0.748
VDA	0.71383	0.095	VDA	0.71164	0.095
LOM	0.13927E-01	0.377	LOM	–0.13266E-01	–0.279
TAA	0.47720E-01	1.459	TAA	0.21520E-01	0.478
VEN	–0.20986E-01	–0.566	VEN	–0.42532E-01	–0.896
FVG	0.24441E-01	0.499	FVG	0.38412E-02	0.067
LIG	0.71383	0.078	LIG	0.72862	0.080
TUS	0.15239E-01	0.370	EMR	–0.36042E-01	–0.718
UMB	0.12074E-01	0.125	TUS	–0.57859E-02	–0.114
MAR	0.14637E-01	0.292	UMB	–0.12084E-01	–0.120
LAZ	–0.79469E-02	–0.145	LAZ	0.24392E-02	0.039
ABR	0.50398E-01	0.904	ABR	0.84764E-01	1.313
MOL	–0.63531E-01	–0.606	MOL	–0.35582E-01	–0.327
CAM	–0.32475E-01	–0.685	CAM	0.53300E-03	0.009
PUG	–0.74071E-01	–1.393	PUG	–0.70844E-01	–1.158
BAS	–0.35821E-02	–0.042	BAS	0.29380E-01	0.324
CAL	–0.14898	–2.550	CAL	–0.10657	–1.579
SIC	–0.15186E-01	–0.308	SIC	–0.35600E-01	0.578
			L(Att)	0.25846E-01	2.403
			SOFF%	–0.62337E-02	–2.580

TABLE 7

A TOBIT ANALYSIS OF THE EFFICIENCY PATTERNS
SPA (223 observations)

Dependent variable = efficiency score from model
Output: Loans, deposits
Input : Employees, branches, free capital

Variable	Coefficient	*t*-ratio	Variable	Coefficient	*t-ratio*
Log-likelihood = –54.82719			*Log-likelihood = –54.78033*		
Constant	1.0413	19.596	Constant	1.0761	7.144
North-West	0.91493E-01	1.503	North-West	0.86818E-01	1.241
North-East	–0.44481E-03	–0.008	North-East	–0.59700E-02	–0.089
Centre	–0.78135E-02	–0.131	Centre	–0.10786E-01	–0.171
South (mainland)	–0.40617E-01	0.653	South (mainland)	–0.37665E-01	0.595
			L(Att)	–0.18132E-02	–0.182
			SOFF%	–0.93476E-03	–0.240
Log-likelihood = –46.26945			*Log-likelihood = –46.20192*		
Constant	1.0573	26.093	Constant	0.74043	5.586
PIE	0.46554E-01	0.685	PIE	0.46392E-01	0.684
VDA	0.90991E-01	1.636	VDA	0.93786E-01	1.671
LOM	–0.10058	–1.241	LOM	–0.10199	–1.258
TAA	–0.12326E-01	–0.189	TAA	–0.11214E-01	–0.172
VEN	–0.31966E-01	–0.417	VEN	–0.34352E-01	–0.447
FVG	–0.60821E-02	–0.062	FVG	–0.54148E-02	–0.055
LIG	0.85261E-01	1.210	LIG	0.84252E-01	1.170
TUS	0.19417E-01	0.226	TUS	0.16453E-01	0.191
UMB	–0.13232	–1.667	UMB	–0.13326	–1.676
MAR	–0.49006E-01	–0.775	MAR	–0.52256E-01	–0.785
LAZ	–0.97845E-01	–1.277	LAZ	–0.99663E-01	–1.240
ABR	0.69404	0.051	ABR	0.68589	0.050
MOL	0.73782E-01	–0.938	MOL	0.66510E-01	0.813
CAM	–0.44624E-01	–0.687	CAM	–0.49633E-01	–0.726
PUG	0.41229E-01	0.318	PUG	0.38501E-01	0.295
BAS	0.12236	1.098	BAS	0.11474	0.972
CAL	–0.42254E-01	–0.655	CAL	–0.49931E-01	–0.671
SIC	0.69404	–0.072	SIC	0.69620	0.072
			L(Att)	–0.36140E-02	–0.357
			SOFF%	0.40044E-03	0.578

6. - Concluding Considerations

This paper has concerned itself with analysing the technical efficiency of a cross-section of 728 Italian banks belonging to different institutional categories (limited companies, credit societies, rural banks and credit co-operatives) using data for 1994 and applying an extension of the non-parametric *FDH* method to take account of input and output slacks. Like the traditional *FDH* approach, this procedure makes no assumptions regarding the functional form of the production technology used by the banks and enables one to allow for the typical multiproduct nature of the banking sector's output.

The main results of the analysis, in which the production set was developed starting from the so-called value added approach, can be summarised as follows. As regards the number of efficient units and the average efficiency score for both the CRCs and the SPAs, the production set whose inputs are employees, branches and free capital would appear to provide the better account of the production performance of the units considered, while the inclusion of securities funds raised among the inputs would not appear to make any significant difference. Another interesting result is that the input structure of the slacks is, in the value added approach, heavily weighted towards capital inputs, showing that the proliferation of branches after 1990 has engendered a significant overbranching in the banking sector.

Moreover, econometric analysis of the determination of the efficiency scores shows that the territorial efficiency patterns differ significantly according to whether CRCs or SPAs are being considered. An attempt was also made to relate the distribution of the efficiency scores to two variables: the size of the bank and the riskiness of the environment (measured by the ratio between bad loans and total loans). These variables are significant only for CRCs, and in this case they diminish the significance of the dummies (of negative sign) for southern Italy. This type of evidence is compatible with the existence of a direct relation between environmental problems in southern Italy and the level of inefficiency of southern Italian CRCs, but in any case requires additional analytical investigation.

BIBLIOGRAPHY

[1] ALY H.Y. - GRABOWSKI R. - PASURKA C. - RANGAN N., «The Technical Efficiency of US Banks», *Economic Letters*, n. 28, 1988, pp. 169-75.

[2] ——————————————, «Technical, Scale and Allocative Efficiencies in US Banking: an Empirical Investigation», *Review of Economic and Statistics*, vol. 27, pp. 211-8.

[3] BANKER R.D. - CHARNES A. - COOPER W.W. - SWARTS J. - THOMAS D.A., «An Introduction to Data Envelopment Analysis with Some of its Models and Their Uses», *Research on Governmental and Nonprofit Accounting*, vol. 5, 1989, pp. 125-63.

[4] BERGER A.N. - HUMPREY D.B., «Measurement and Efficiency Issues in Commercial Banking», FED Finance and Economics, *Discussion Series*, n. 151, 1990.

[5] CANNARI L. - SIGNORINI L.F., *Rischiosità e razionamento: un'analisi dell'efficienza allocativa delle banche di credito cooperative*, Roma, Banca d'Italia, mimeo, 1995.

[6] CHARNES A. - COOPER W.W. - RHODES E., «Measuring the Efficiency of Decision-Making Units», *European Journal of Operational Research*, vol. 2, 1978, pp. 429-44.

[7] DEBREU G., «The Coefficient of Resource Utilisation», *Econometrica*, vol. 19, 1951, pp. 273-92.

[8] DEPRINS D. - SIMAR L. - TULKENS H., «Measuring Labour Efficiency in Post Offices» in MARCHAND M. - PESTIEAU P. - TULKENS H. (eds.), *The Performance of Public Enterprises: Concepts and Measurement*, Amsterdam, North Holland, 1984, pp. 243-67.

[9] DESTEFANIS S. - PAVONE A., «L'analisi dell'efficienza nell'ambito dell'approccio FDH. Un'estensione e un'applicazione per gli ospedali di zona», *Quaderni di Ricerca*, n. 3, Roma, ISTAT, 1996.

[10] DRAKE L. - WEYMAN-JONES T.G., «Technical and Scale Efficiency in UK Building Societies», *Applied Financial Economics*, vol. 2, 1992, pp. 1-9.

[11] ELYASIANI E. - MEHDIAN S., «Productive Efficiency Performance of Minority and Nonminority-Owned Banks: A Nonparametric Approach», *Journal of Banking and Finance*, vol. 16, 1992, pp. 933-48.

[12] FARRELL M.J., «The Measurement of Productive Efficiency», *Journal of the Royal Statistical Society*, Series A, General, vol. 120, 1957, pp. 253-81.

[13] FERRIER G.D. - LOVELL C.A.K., «Measuring Cost-Efficiency in Banking», *Journal of Econometrics*, vol. 46, 1990, pp. 229-45.

[14] FRIED H.O. - LOVELL C.A.K. - SCHMIDT S. (eds), *The Measurement of Productive Efficiency: Techniques and Applications*, London, Oxford University Press, 1993.

[15] FRIED H.O. - LOVELL C.A.K. - VANDEN ECKAUT P., «Evaluating the Performance of US Credit Unions», *Journal of Banking and Finance*, vol. 17, 1993, pp. 251-65.

[16] KOOPMANS T.C., «An Analysis of Production as an Efficient Combination of Activities», in KOOPMANS T.C. (ed.), *Activity Analysis of Production and Allocation*, Chapter 3, New York, Wiley, 1951.

[17] ——, *Three Essays on the State of Economic Science*, New York, McGraw Hill, 1957.

[18] LOCATELLI R. - PROSPERETTI L. (eds.), *La produttività nelle aziende di credito - analisi su dati di bilancio*, Milano, Sole 24 Ore, 1987.

[19] LOVELL C.A.K., «Production Frontiers and Productive Efficiency», in FRIED H.O. - LOVELL C.A.K. - SCHMIDT S. (eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, London Oxford University Press, 1993.

[20] ONADO M., *Economia dei sistemi finanziari*, Bologna, il Mulino, 1992.

[21] RESTI A., «Il dibattito su efficienza delle banche e economie di scala: il contributo della Data Envelopment Analysis ed un'applicazione a dati italiani», *Politica Economica*, vol. 10, 1994, pp. 269-311.

[22] THIRY B. - TULKENS H., «Allowing for Technical Inefficiency in Parametric Estimation of Production Functions for Urban Transit Firms», *Journal of Productivity Analysis*, vol. 3, 1992, pp. 45-65.

[23] TULKENS H., «La performance productive d'un service public. Définition, méthodes de mesure et application à la Régie des Postes en Belgique», *L'Actualité Economique, Revue d'Analyse Economique*, Montreal, vol. 62, 1986; pp. 306-35.

[24] ——, «Non-Parametric Efficiency Analyses in Four Service Activities: Retail Banking, Municipalities, Courts and Urban Transit», Louvain-la-Neuve, Université Catholique de Louvain, *Discussion Paper*, n. 9050, 1990.

[25] ——, «On FDH Efficiency Analyses: Some Methodological Issues and Applications to Retail Banking, Municipalities, Courts and Urban Transit», *Journal of Productivity Analysis*, vol. 4, 1993, pp. 179-210.