ISSN 1748-7595 (Online)

Kent Business School

Working Paper Series

Disserting Banking Dea Efficiency

C. Serrano Cinca
C. Mar. Molinero
Y. Fuertes Callen
Kent Business School

KBS
UNIVERSITY OF KENT

Working Paper No. 92
Jan 2006

DISSECTING BANKING DEA EFFICIENCY

C. Serrano Cinca*
Department of Accounting and Finance
University of Saragossa, Spain.

C. Mar Molinero
Kent Business School
University of Kent, UK.

Y. Fuertes Callén
Department of Accounting and Finance
University of Saragossa, Spain.

ABSTRACT

DEA has been extensively used to measure the efficiency of financial institutions. The advantages of this approach are clearly understood but many problems remain unsolved. One of them is the selection of inputs and outputs. There are various views of what constitutes inputs and outputs in a financial institution. The paper explores up to what point the various combinations of inputs and outputs are equivalent, and up to what point the efficiency score obtained by a given institution changes under the various combinations of inputs and outputs. The extent to which two institutions that achieve the same efficiency score arrive at it following different strategies is explored with the aim of finding out what is behind such a score. By-products of the approach proposed here are the creation of league tables of financial institutions in terms of efficiencies and the possibility of assessing the strengths and weaknesses of individual institutions. This methodology is applied to the particular case of American banks efficiency.

KEY WORDS: Efficiency, Principal Component Analysis, banking, Data Envelopment Analysis.

JEL: G21

* ***Carlos Serrano-Cinca:*** Department of Accounting and Finance, Fac. CC Económicas y Empresariales, Univ. Zaragoza, Gran Vía 2, Zaragoza (50.005) SPAIN serrano@unizar.es http://ciberconta.unizar.es/charles.htm

DISECTING BANKING DEA EFFICIENCY

1. Introduction

Efficiency is a key concept for financial institutions, and it has long been studied. A review of 130 such studies in 21 countries is given by Berger and Humphrey (1997). Berger and Humphrey classify papers according to the technical approach employed, which they identify as parametric- Stochastic Frontier Approach (SFA), Distribution Free Approach (DFA), Thick Frontier Approach (TFA)- or non parametric- Data Envelopment Analysis (DEA), Free Disposal Hull (FDH), Index Numbers (IN), Mixed Optimal Strategy (MOS). By far the most popular technical approach is DEA, which was applied in 62 of the papers surveyed.

DEA is becoming widely used to assess the efficiency of organizations with multiple homogeneous decision units that produce several outputs with a variety of inputs. For an extensive bibliography of DEA see, for example, Emrouznejad and Thanassoulis (1996), and Seiford (1996). DEA is appropriate for sets of homogeneous units with similar inputs and similar outputs since it performs multiple comparisons using a Linear Programming based approach. The assumptions underlying DEA are minimal. Inputs and outputs can be measured in their own units, and these units can be different for the different inputs and outputs. A survey of the more restricted area of DEA applications to bank branch performance is given by Schaffnit et al (1997). Some recent references on the application of DEA to financial institutions are Athanassopoulos (1997), Pastor et al (1997), Seiford and Zhu (1999), Saha and Ravisankar (2000), Dekker and Post (2001), Kuosmanen and Post (2001), Hartman et al (2001), and Luo (2003).

For the purposes of this paper, it will be useful to make a distinction between model and specification in a DEA context. Different philosophical approaches as to what a financial institution does, and what is meant by efficiency will lead to different models. Two basic models are prevalent in the literature: intermediation and production; see Berger and Mester (1997) for a full discussion. Specification will refer to a more restricted concept: the particular set of inputs and outputs that enter into model definition.

The variety of models and specifications for financial efficiency analysis is reflected in practice. The selection of inputs and outputs varies from study to study, giving an impression of confusion. For example; a particular item, such as deposits, may be treated as an input or as an output according to whether the institution is modelled from the point of view of production or from the point of view of intermediation, see Athanassoupoulos (1997). This is a matter of concern, as the level of efficiency of a financial institution may depend on the particular choice of inputs and outputs. It may be puzzling for the manager of a bank branch to discover that it is possible for different researchers to arrive at different conclusions about the efficiency of a bank branch when using the same technique (DEA). However, this confusion may be more apparent than real, since alternative specifications may be equivalent. The study of the extent to which two different specifications are equivalent is one of the purposes of this paper.

Model and specification selection are not the only issues addressed in this paper. We wish to go behind the efficiency score. Two financial institutions may achieve the same DEA efficiency under a given model and under a common specification, but they may still be very different. Efficiency, being a mere score, may be compatible with a variety of management strategies. Imagine two institutions that achieve the same efficiency, one may have specialised in the production of a particular output and the other on the good use of a particular input. These differences will, of course, be reflected in different weight structures for inputs and outputs, and could be identified by means of such techniques as cross-efficiency analysis; Doyle and Green (1994). Here we apply a recent methodological approach based on DEA and multivariate statistical analysis; Serrano-Cinca and Mar Molinero (2004), and Serrano Cinca et al (2005). This approach has the advantage of visualising the way in which a particular DEA score has been achieved by a financial institution, and how this score is related to the model selected.

In this paper, efficiencies are calculated for a variety of DEA specifications. It is proposed that DEA modelling be embedded in a multivariate statistical framework.

This paper unfolds as follows. The next section contains a discussion of efficiency in financial institutions. The particular case study of American banks efficiency is introduced

and presented in the next section. This is followed by a description of the model and its implementation. The paper is completed with a conclusions section.

2. Efficiency modelling in financial institutions

For modelling purposes, financial institutions are seen from the point of view of intermediation or from the point of view of production; see Athanassoupoulos (1997). Under the intermediation model they collect deposits and make loans in order to make a profit. Deposits and acquired loans are inputs. Institutions are interested in placing loans, which are traditional outputs in studies of this kind; see, for example Berger and Humphrey (1991). Under the production model, a financial institution uses physical resources such as labour and plant in order to process transactions, take deposits, lend funds, and so on. In the production model manpower and assets are treated as inputs and transactions dealt with -such as deposits and loans- are treated as outputs. See, for example, Vassiloglou and Giokas (1990), Schaffnit et al (1997), Soteriou and Zenios (1999).

The mathematical models used to study the efficiency of financial institutions can be divided into two groups: those based on parametric frontier techniques, and those based on Data Envelopment Analysis (DEA). Berger and Humphrey (1991) find inconsistencies between the two approaches, although Ondrich and Ruggiero (2001) argue that both produce similar rankings, and conclude that there is no advantage in using parametric frontiers.

In this paper we focus on DEA models. Up to what point different DEA modelling approaches produce different results? This question can only be answered by looking at particular case studies. Oral and Yolalan (1990) found that a DEA model aimed at estimating service efficiency in bank branches in Turkey produced indistinguishable results from an alternative DEA model focused on profitability. A possible way out would be to develop specifications with many inputs and outputs. This would be an attempt to create a general model that encompasses various modelling philosophies as particular cases. But care has to be exercised since the more inputs and outputs a model contains, the more units become efficient through specialisation or, as Lovell and Pastor (1997) put it, "because they are self-identifiers". The relationship between efficiency and the number of inputs and outputs has been studied by Pedraja Chaparro et al (1999).

Alternative specifications for inputs and outputs for a given model have been explored in many studies. Athanassopoulos (1997) observes a lack of consistency in the selection of inputs and outputs when studying bank branch efficiency. Oral and Yolalan (1990) experiment with various specifications and observe that efficiencies change according to the input/output mix chosen. Some times there is no choice, as the chosen specification is in part determined by the data that is available; Vassiloglou and Giokas (1990). Lovell and Pastor (1997) observe that alternative specifications may not give significantly different results, and apply the Pastor et al (2002) methodology to choose a parsimonious specification. This approach is based on a sound mathematical model, but has a mechanical feel to it. But different specifications are not totally equivalent, and it is difficult to assess what are the consequences for individual units of adding or removing an input/output without engaging in considerable extra work.

A new approach to specification search is applied in this paper; Serrano-Cinca and Mar Molinero (2004), and Serrano Cinca et al (2005). The distinctive features of a specification are revealed by embedding DEA efficiency results into a multivariate statistical framework. We use in particular Principal Components Analysis (PCA), multiple regression, and Hierarchical Cluster Analysis (HCA). PCA has been used as an alternative to DEA by Zhu (1998) and Premachandra (2001). PCA as a data reduction technique to select inputs and outputs has been used by Adler and Golany (2001).

In this approach, PCA plays a fundamental role in specification and model selection. We do not attempt to find a "best" specification of inputs and outputs. A variety of possible specifications that offer combinations of inputs and outputs are estimated and efficiencies calculated for each financial institution under each specification. In this way, a matrix is obtained in which each column corresponds to a specification, and each row to a financial institution. This matrix is analysed by means of Principal Components Analysis (PCA). Component scores are plotted to show the extent to which the efficiency of financial institutions remains unchanged under the various specifications. The plot is interpreted by means of property fitting (Pro-Fit), a regression-based technique. The superimposition of the Pro-Fit results on the scores plot will help to identify specification equivalence, guide model

selection, identify outlying behaviour, and assess strategic behaviour patterns in financial institutions that achieve the same efficiency score.

3. A case study: American banks

US commercial banks are by far the best studied financial institutions from the point of view of efficiency. Their study has been undertaken from a variety of perspectives and using several methodologies: Stochastic Frontier Approach (SFA), Distribution Free Approach (DFA), Thick Frontier Approach (TFA), Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH); see Berger and Humphrey (1997). Amongst the studies that have applied DEA to the analysis of efficiency in US banks we can list Aly et al (1990), Barr et al (1993), Berg et al (1992 and 1993), Charnes et al (1990), Elyasiani and Mehdian (1995), Ferrier and Lovell (1990), Thompson et al (1997), and Miller and Noulas (1996).

According to Barr et al (2002), the environment in which U.S. commercial banks operate has undertaken profound transformations, as has the way in which they conduct their business. They mention changes in the regulatory environment, in the impact of information technology, and in the way in which they assess risks. A very competitive industry has emerged. They further point out that such changes tend to result in a split between institutions that "perform relatively well and those that perform relatively poorly", and that the different kinds of institutions can be identified using non-parametric methods.

This section will be divided into sub-sections. First, the data set will be described. The second subheading will concentrate on DEA and PCA. Empirical results will be interpreted in the third and fourth sub-sections.

3.1. The data set: 3 inputs and 3 outputs

Data was obtained from Standard & Poor's COMPUSTAT database for the year 2003, SIC code 6020, "Commercial Banks". All of them are incorporated in the US and are quoted in the New York stock exchange -or in Nasdaq-. This ensures that the sample is homogeneous, a condition for the application of DEA. In total, 85 American banks met the conditions and all of them were selected. Having been extracted from annual accounts, all the data except

number of employees, is measured in monetary units. The list of all institutions is given in Table 1. Rather than use the full name of each institution, the "trading symbol" has been employed.

Trading Symbol	*Labour (number of employees)*	*Physical capital ($ millions)*	*Deposits ($ millions)*	*Interest and non-interest income ($ millions)*	*Loans ($ millions)*
ASO	12,385	964.692	30,440.352	2,942.229	29,057.531
ASBC	4,091	131.315	9,792.840	972.499	10,225.957
BAC	133,549	6,036.000	414,113.000	49,006.000	365,300.000
BOH	2,700	160.005	7,332.777	641.241	5,792.040
BK	22,901	1,398.000	56,406.000	6,336.000	42,901.000
BNK	6,700	264.818	17,901.184	1,560.128	16,155.371
BBT	26,300	1,201.342	59,349.785	6,243.924	61,791.606
BOKF	3,449	175.901	9,219.863	868.165	7,355.250
BPFH	437	13.740	1,658.461	158.764	1,597.292
CHZ	2,058	75.179	4,969.891	371.543	3,692.482
CIH	212,400	6,514.000	478,494.000	64,120.000	495,332.000
CBCF	2,342	112.784	5,442.266	500.796	5,166.832
CYN	2,348	62.719	10,937.063	752.950	8,122.501
CNB	3,939	246.170	9,768.590	908.257	11,845.233
CMA	11,282	374.000	41,463.000	3,299.000	39,499.000
CBH	8,200	811.451	20,701.398	1,248.109	7,371.285
CBSH	4,967	336.366	10,206.207	919.077	8,007.457
CBU	1,259	61.705	2,725.488	228.710	2,099.414
CBSS	7,700	527.295	15,687.820	1,801.343	17,120.918
CORS	468	26.313	2,846.402	187.159	2,397.323
CFR	3,268	168.611	8,068.857	584.307	4,510.035
EWBC	730	24.957	3,312.667	211.322	3,234.133
FNB	1,682	199.735	6,159.496	553.883	5,657.201
FITB	18,899	1,828.000	57,095.000	6,474.000	53,419.000
FBP	1,983	85.269	6,765.105	624.507	7,624.077
FCTR	1,031	95.756	2,427.897	240.210	2,255.798
FCF	1,474	46.538	3,288.275	289.176	2,792.859
FHN	11,494	350.202	15,679.969	2,670.886	16,808.412
FMBI	1,646	91.535	4,815.105	365.427	4,012.998
FMER	3,063	119.079	7,502.781	777.415	6,517.363
FULT	2,950	120.777	6,751.781	572.518	6,115.055
GBBK	1,710	83.816	5,312.664	579.261	4,411.637
HNBC	623	23.329	1,979.081	146.838	1,400.189
HIB	5,339	217.399	14,159.516	1,260.388	12,878.136
HU	1,859	125.168	6,243.355	527.174	4,633.264
HBAN	7,983	349.712	18,487.395	2,361.797	22,260.658
IBNK	886	54.563	1,812.630	174.852	1,674.527
IFIN	2,413	76.420	4,207.117	582.256	199.530
IFC	3,589	32.208	2,899.662	700.283	3,980.664
JPM	93,453	6,487.000	326,492.000	44,363.000	225,170.000
KEY	20,034	606.000	50,858.000	5,730.000	61,305.000
LBAI	502	27.510	1,325.682	79.705	834.637
MTB	14,000	398.971	33,114.945	2,957.660	35,171.573
MI	12,244	438.485	22,270.105	2,740.721	24,835.379

KRB	26,500	2,676.597	31,836.078	11,684.361	35,998.091
MEL	20,900	668.000	20,843.000	4,403.000	10,139.000
MRBK	3,565	140.922	10,262.551	773.166	9,146.329
NARA	320	6.766	1,061.415	81.802	988.795
NCC	33,331	1,125.526	63,930.031	9,593.821	93,521.063
NPBC	1,074	43.653	2,435.296	206.933	2,223.667
NBY	477	14.768	815.839	72.222	582.933
NFB	2,979	150.875	15,116.113	1,255.091	12,222.539
NTRS	8,056	498.300	26,270.000	2,580.100	26,432.297
ONB	3,019	181.398	6,493.090	661.897	5,496.387
PRK	1,645	36.746	3,414.249	320.152	2,667.661
PNC	23,200	1,456.000	45,241.000	5,969.000	34,848.000
PVTB	219	6.233	1,547.359	101.442	1,213.977
PBKS	1,629	49.575	3,079.549	343.790	2,754.023
PVN	4,525	84.198	10,101.055	2,781.408	5,655.070
RF	16,180	629.638	32,732.535	3,617.887	33,068.654
RIGS	1,450	226.502	4,286.230	346.932	3,483.946
SIVB	969	14.999	3,666.876	277.397	1,924.729
SKYF	3,546	154.242	8,514.852	844.285	9,361.842
TSFG	1,918	142.705	6,028.648	509.618	5,690.583
SWBT	1,760	117.951	4,403.238	326.023	3,545.564
STT	19,850	1,212.000	47,516.000	4,727.000	26,698.000
SBIB	1,036	48.541	2,418.369	200.944	2,127.675
STI	27,578	1,595.307	81,189.500	7,071.841	85,358.766
SUSQ	2,065	62.961	4,134.465	387.770	4,220.598
SNV	10,909	791.439	15,941.609	2,430.821	16,376.583
TCB	8,136	282.193	7,611.746	1,105.143	8,606.531
TRST	488	20.168	2,419.810	166.779	1,113.527
TRMK	2,356	108.374	5,089.457	516.931	4,958.336
USB	51,377	1,957.000	11,9052.000	14,571.000	117,299.000
UCBH	666	84.145	4,483.520	283.465	3,730.780
UB	10,146	509.734	35,532.281	2,563.916	25,636.939
UBSI	1,585	46.354	4,182.371	400.824	4,264.482
VLY	2,264	128.606	7,162.965	605.695	6,107.758
WB	86,670	4,619.000	221,225.000	24,474.000	165,375.000
WFC	140,000	3,534.000	247,527.000	31,800.000	285,706.000
WABC	1,003	35.748	3,463.991	268.575	2,269.420
WTNY	2,369	148.259	6,158.582	427.573	4,838.441
WL	2,307	152.300	6,577.199	633.000	6,135.398
WTFC	929	156.714	3,876.621	276.083	3,302.522
ZION	7,896	407.825	20,896.695	1,889.483	19,652.739

Table 1: List of banks and the values of inputs and outputs. Physical capital, Deposits, Interest and non-interest income, and Loans in millions of U.S. dollars. Number of employees in units.

There is much agreement on what constitutes inputs and outputs under the production model and under the intermediation model, although not all authors use the same set of inputs and outputs. After a thorough survey of the inputs and outputs used in the literature, -see table 2-

the following inputs and outputs were selected. Inputs: labour, physical capital, and deposits. Outputs: interest and non-interest income; deposits; and loans.

Variable symbol	*Variable Name*	*Description and COMPUSTAT acronimous*	*Used by*
Input A	Number of employees	Employees "EMP"	Sherman and Gold (1985), Ferrier and Lovell (1990), Aly et al (1990), Berg et al (1993), Seiford and Zhu (1999), Wheelock and Wilson (1999), and Luo (2003)
Input B	Physical capital	Fixed Assets (Net) "PPENT"	Berg et al (1993), Seiford and Zhu (1999), Wheelock and Wilson (1999), Tortosa-Ausina (2002), Luo (2003), Barr et al (2002), Aly et al (1990), and Berg et al (1991)
Input C	Deposits	Deposits Customer "DPTC" + Deposits Banks "DPTB"	Sealey and Lindley (1977), Mester (1989), Elyasiani and Mehdian (1992), Miller and Noulas (1996), Mester (1997), Brockett et al (1997), and Casu and Girardone (2004)
Output 1	Interest and non-interest income	Interest & Div Inc Total "IDIT" + Income Noninterest Tot Bank "INITB"	Miller and Noulas (1996), Thompson et al (1997), Brockett et al (1997) and Seiford and Zhu (1999)
Output 2	Deposits	Deposits Customer "DPTC" + Deposits Banks "DPTB"	Sherman and Gold (1985), Ferrier and Lovell (1990), Berger and Humphrey (1991). Berg et al (1991), Berg et al (1993), and Kumbhakar et al (2001)
Output 3	Loans	Loans/Claims/Advances Banks & Govt "LCABG" + Loans/Claims/Advances Customers "LCACU"	Sherman and Gold (1985), Ferrier and Lovell (1990), Aly et al (1990), Berger and Humphrey (1991), Berg et al (1991), Berg et al (1993), English et al (1993), Miller and Noulas (1996), Mester (1997), Wheelock and Wilson (1999), Brockett et al (1997), and Casu and Girardone (2004)

Table 2: Inputs and outputs

Labour

This item measures the number of company workers as reported to shareholders. Labour has been used in banking efficiency studies by Sherman and Gold (1985), Ferrier and Lovell (1990), Aly et al (1990), Berg et al (1993), Seiford and Zhu (1999), Wheelock and Wilson (1999), and Luo (2003).

Physical capital

This item measures the net cost or valuation of tangible fixed assets used in the regular business operations of the company, less accumulated depreciation, investment grants, and other deductions. It has been used by Berg et al (1993), Seiford and Zhu (1999), Wheelock and Wilson (1999), Tortosa-Ausina (2002), Luo (2003), Barr et al (2002), Aly et al (1990) and Berg et al (1991).

Deposits

This item measures the total demand, savings, and time deposits held on account for individuals, partnerships, and corporations plus deposits held on account for other banks. It is an issue whether deposits are inputs or outputs. See Pastor et al (1997) for a discussion. Deposits are treated as inputs by Sealey and Lindley (1977), Mester (1989), Elyasiani and Mehdian (1992), Miller and Noulas (1996), Mester (1997), Brockett et al (1997), and Casu and Girardone (2004); they are treated as outputs by Sherman and Gold (1985), Ferrier and Lovell (1990), Berger and Humphrey (1991). Berg et al (1991), Berg et al. (1993), and Kumbhakar et al (2001); they are treated simultaneously as inputs and outputs by Aly et al (1990), Lozano (1998), Wheelock and Wilson (1999) and Tortosa-Ausina (2002).

Interest and non-interest income

This item measures revenue received from all earning assets plus total revenue/income that cannot be attributed to interest and dividends received from earning assets. Used by Miller and Noulas (1996), Thompson et al (1997), Brockett et al (1997) and Seiford and Zhu (1999).

Loans

This item measures the monetary value of all outstanding loans, claims, and advances made to individual, commercial, and industrial borrowers, less reserves for possible credit losses and unearned income. Used by Sherman and Gold (1985), Ferrier and Lovell (1990), Aly et al (1990), Berger and Humphrey (1991), Berg et al (1991), Berg et al (1993), English et al (1993), Miller and Noulas (1996), Mester (1997), Wheelock and Wilson (1999), Brockett et al (1997) and Casu and Girardone (2004).

The values of all inputs and outputs for all the banks are given in Table 1.

Notation will be introduced in order to simplify the discussion of the various specifications. Inputs are referred to by means of capital letters, in such a way that the first input is represented by the letter A, the second input by the letter B, and the third one by the letter C. Outputs are referred to by means of numbers. The first input is associated with number 1, the second input with number 2, and the third input with number 3. In this way a specification that treats a bank as an institution whose employees (input A) take deposits (output 2) and place loans in the market (output 3) would be labeled A23. If this specification is augmented with physical assets (input B) and income (output 1), the specification becomes AB123. Specification AB123 treats a bank as a production unit that employs manpower (A) and plant (B) in order to generate income, deposits, and loans. An intermediation model would be described by a specification such as AC13, in which deposits (C) are treated as an input. Under the AC13 specification a bank is an institution whose employees collect deposits in order to make loans and generate income.

Other possible views of the way in which a bank operates can be generated by using different combinations of inputs and outputs. Efficiency ratios are generated by choosing a specification with only one output and one input. It is, of course, possible to use all possible combinations of inputs with all possible combinations of outputs. The total number of specifications that could possibly be generated with n inputs and m outputs is given by the formula

$$\sum_{i=1}^{n} C_n^i * \sum_{i=1}^{m} C_m^i \quad \text{where } C_n^i = \binom{n}{i} = \frac{n!}{i!(n-i)!}$$

In general, it will not be necessary to calculate efficiencies under all possible specifications, as some of them can be discarded on a priori grounds. In our case there are 3 inputs and 3 outputs, giving a possible total number of specifications of 49. Specifications that treat deposits both as inputs and outputs have been excluded, reducing their total number to 33. The complete list of specifications and the inputs and outputs that they contain can be found in Table 3.

Model	*INPUT*	*OUTPUT*
a1	Employees	Income
a12	Employees	Income, Deposits
a123	Employees	Income, Deposits, Loans
a13	Employees	Income, Loans
a23	Employees	Deposits, Loans
a2	Employees	Deposits
a3	Employees	Loans
b1	Physical Assets	Income
b12	Physical Assets	Income, Deposits
b123	Physical Assets	Income, Deposits, Loans
b13	Physical Assets	Income, Loans
b23	Physical Assets	Deposits, Loans
b2	Physical Assets	Deposits
b3	Physical Assets	Loans
ab1	Employees, Physical Assets	Income
ab12	Employees, Physical Assets	Income, Deposits
ab123	Employees, Physical Assets	Income, Deposits, Loans
ab13	Employees, Physical Assets	Income, Loans
ab23	Employees, Physical Assets	Deposits, Loans
ab2	Employees, Physical Assets	Deposits
ab3	Employees, Physical Assets	Loans
c1	Deposits	Income
c13	Deposits	Income, Loans
c3	Deposits	Loans
ac1	Employees, Deposits	Income
ac13	Employees, Deposits	Income, Loans
ac3	Employees, Deposits	Loans
bc1	Physical Assets, Deposits	Income
bc13	Physical Assets, Deposits	Income, Loans
bc3	Physical Assets, Deposits	Loans
abc1	Employees, Physical Assets, Deposits	Income
abc13	Employees, Physical Assets, Deposits	Income, Loans
abc3	Employees, Physical Assets, Deposits	Loans

Table 3: The 33 specifications and their definitions

DEA efficiencies, on a scale from 0% to 100%, for all banks were calculated under Variable Returns to Scale (VRS) for all specifications. The results are given in Table 4.

	a1	a12	a123	a13	a23	a2	a3	b1	b12	b123	b13	b23	b2	b3	ab1	ab12	ab123	ab13	ab23	ab2	ab3	c1	c13	c3	ac1	ac13	ac3	bc1	bc13	bc3	abc1	abc13	abc3
ASO	39	60	69	69	65	60	65	10	28	29	29	28	27	28	25	49	51	51	47	46	47	28	66	66	37	75	74	24	50	50	24	50	50
ASBC	40	51	60	60	60	45	60	24	45	64	64	64	42	64	32	52	69	69	68	47	68	33	73	73	42	80	79	33	78	78	33	78	78
BAC	91	100	100	100	100	100	100	82	93	93	82	93	93	77	99	100	100	100	100	100	97	84	85	81	89	100	99	85	85	79	85	85	79
BOH	40	52	52	52	48	48	45	14	25	28	28	27	24	27	27	45	47	47	43	41	43	32	56	56	42	69	63	27	45	45	27	45	45
BK	53	65	66	65	66	65	55	33	44	44	37	42	42	29	49	62	62	57	57	57	42	31	55	52	46	68	59	49	53	44	49	53	44
BNK	38	57	62	62	62	57	62	18	52	55	55	54	51	54	28	61	64	64	61	58	61	27	63	63	36	73	73	26	65	65	26	65	65
BBT	45	60	74	74	73	60	73	37	54	55	53	55	52	51	48	63	67	67	67	59	64	29	71	71	40	77	77	49	66	66	49	66	66
BOKF	42	55	55	55	49	49	48	16	31	34	34	32	29	32	29	48	51	51	46	43	46	32	56	56	42	70	63	28	49	49	28	49	49
BPFH	71	71	76	76	66	54	66	58	58	70	70	68	49	68	71	71	79	79	72	55	72	63	87	87	88	99	93	72	87	87	88	99	93
CHZ	32	40	40	39	38	38	32	19	31	36	36	35	31	35	32	47	47	45	46	46	41	33	55	55	39	60	58	30	47	47	39	60	58
CIH	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CBCF	37	45	48	48	45	37	45	16	25	34	34	34	24	34	26	41	49	49	47	37	47	36	68	68	43	73	72	31	55	55	31	55	55
CYN	54	88	88	79	88	88	79	40	100	100	100	100	100	100	54	100	100	100	100	100	100	24	52	52	44	81	79	36	88	88	44	92	92
CNB	38	51	73	73	73	46	73	12	24	40	40	40	22	40	25	42	63	63	63	38	63	31	85	85	40	94	94	26	68	68	26	68	68
CMA	50	94	100	100	100	94	100	37	100	100	100	100	100	100	46	100	100	100	100	100	100	23	65	65	39	94	94	35	91	91	35	91	91
CBH	25	56	56	30	56	56	20	5	20	20	8	20	20	7	15	41	41	21	41	41	16	19	26	25	25	35	28	14	19	17	14	19	17
CBSH	31	41	41	41	38	38	37	9	18	19	19	18	17	18	20	34	35	35	32	31	32	30	55	55	35	59	56	24	40	40	24	40	40
CBU	34	38	38	37	32	32	30	16	19	23	23	22	18	22	34	39	39	38	36	36	33	45	64	64	50	63	60	37	45	45	50	63	60
CBSS	38	49	60	60	58	41	58	11	23	30	30	29	21	29	24	40	49	49	46	35	46	35	76	76	41	78	77	27	56	56	27	56	56
CORS	76	91	92	92	92	89	92	33	44	60	60	60	44	60	76	99	100	100	100	99	100	39	68	68	70	97	97	43	69	69	70	100	99
CFR	30	44	44	36	44	44	27	12	26	26	21	26	26	20	21	40	40	33	39	39	28	27	40	40	34	49	43	23	32	32	23	32	32
EWBC	54	67	79	79	79	67	79	38	54	91	91	91	54	91	54	82	100	100	100	82	100	36	74	74	56	99	99	44	89	89	56	100	100
FNB	56	71	75	75	71	61	71	10	16	21	21	21	16	21	33	51	56	56	54	46	54	34	66	66	52	90	85	27	46	46	27	46	46
FITB	66	80	93	93	87	80	86	26	35	35	32	32	32	29	51	65	67	67	60	60	59	31	65	64	53	88	82	46	52	49	46	52	49
FBP	53	68	87	87	87	59	87	25	44	69	69	69	41	69	53	70	95	95	95	66	95	34	80	80	50	100	100	34	84	84	50	100	100
FCTR	43	46	48	48	39	34	39	11	12	16	16	15	10	15	43	46	48	48	39	34	39	52	76	76	60	77	74	38	47	47	60	77	74
FCF	35	40	41	41	34	33	34	25	31	42	42	41	29	41	35	45	49	49	44	41	44	42	66	66	49	67	65	40	57	57	49	68	66
FHN	38	43	49	49	38	27	38	23	38	50	50	43	32	43	30	43	53	53	42	32	42	49	78	74	50	73	64	43	73	64	43	73	64
FMBI	39	49	50	50	45	45	45	15	25	32	32	32	25	32	39	52	53	53	53	52	53	33	61	61	44	72	70	29	48	48	44	72	70
FMER	43	53	55	55	47	43	47	22	36	43	43	42	34	42	32	50	56	56	51	43	51	36	62	62	46	73	66	34	57	57	34	57	57

FULT	33	43	45	45	45	39	45	17	30	38	38	38	29	38	25	42	49	49	48	39	48	32	64	64	38	69	68	29	56	56	29	56	56
GBBK	58	67	68	68	50	50	50	24	34	42	41	39	31	39	58	69	71	71	59	57	58	41	60	60	58	83	71	38	52	52	58	83	71
HNBC	47	50	50	48	46	46	40	32	35	36	35	35	34	33	47	51	51	48	49	49	41	51	66	66	66	69	65	51	55	55	66	69	65
HIB	39	54	60	60	60	52	60	18	46	52	52	51	44	51	29	56	62	62	60	53	60	28	63	63	38	74	73	27	64	64	27	64	64
HU	49	62	62	60	56	56	49	15	26	29	28	27	25	27	32	52	52	52	48	48	46	33	54	54	48	74	66	28	43	43	28	43	43
HBAN	48	60	80	80	76	50	76	21	44	61	61	58	40	58	34	58	74	74	69	49	69	38	83	83	48	92	91	34	78	78	34	78	78
IBNK	38	38	41	41	34	29	34	15	15	19	19	18	13	18	38	38	41	41	34	29	34	60	82	82	64	74	73	47	52	52	64	74	73
IFIN	41	44	44	41	26	26	9	26	30	30	26	24	24	8	41	45	45	41	33	33	9	52	52	19	55	55	19	46	46	17	55	55	19
IFC	33	33	36	36	20	12	20	73	73	100	100	90	37	90	73	73	100	100	90	37	90	86	100	100	69	85	73	92	100	100	92	100	100
JPM	100	100	100	100	100	100	86	68	68	68	68	68	68	42	100	100	100	100	92	92	71	95	95	58	100	100	80	84	84	50	84	84	50
KEY	54	67	96	96	95	67	95	65	91	100	100	100	84	100	69	88	100	100	100	78	100	31	83	83	46	96	96	65	100	100	65	100	100
LBAI	44	44	44	44	44	44	44	23	23	23	23	23	23	23	44	44	44	44	44	44	44	63	73	73	72	72	72	52	52	52	72	72	72
MTB	35	59	71	71	71	59	71	26	74	83	83	83	72	83	31	70	77	77	77	68	77	26	73	73	34	81	81	30	87	87	30	87	87
MI	36	47	60	60	56	41	56	19	44	55	55	52	41	52	28	50	60	60	55	43	55	36	77	77	41	77	75	32	71	71	32	71	71
KRB	89	89	89	89	39	30	39	38	38	38	38	13	10	13	72	72	72	72	26	21	26	100	100	78	100	100	63	100	100	47	100	100	47
MEL	38	38	38	38	22	22	12	38	38	38	38	25	25	12	43	43	43	43	25	25	13	59	59	34	54	54	26	73	73	25	73	73	25
MRBK	36	54	60	60	60	54	60	18	42	52	52	52	41	52	27	53	63	63	63	52	63	26	63	63	36	74	74	26	65	65	26	65	65
NARA	68	68	68	68	68	68	68	92	92	92	92	92	92	92	92	92	92	92	92	92	92	79	100	100	100	100	100	100	100	100	100	100	100
NCC	57	58	93	93	92	51	92	71	72	92	92	91	61	91	77	77	98	98	98	60	98	41	100	100	53	100	100	84	100	100	84	100	100
NPBC	36	40	42	42	37	33	37	21	24	33	33	33	23	33	36	41	44	44	43	37	43	48	75	75	54	73	72	42	57	57	54	73	72
NBY	46	46	46	46	46	46	46	42	42	42	42	42	42	42	46	46	46	46	46	46	46	100	100	100	99	99	99	88	88	88	99	99	99
NFB	70	100	100	100	100	100	100	26	71	71	71	70	70	69	48	99	99	98	95	95	93	26	56	56	52	98	94	30	70	70	30	70	70
NTRS	52	77	94	94	91	77	91	16	46	50	50	49	44	49	34	69	75	75	71	64	71	29	69	69	44	92	92	26	65	65	26	65	65
ONB	37	45	47	47	38	37	38	12	20	24	24	23	19	23	25	37	41	41	36	32	36	37	61	61	43	67	62	30	44	44	30	44	44
PRK	35	39	39	39	31	31	29	34	42	53	52	49	38	49	35	48	53	53	49	44	49	43	61	61	48	63	59	46	59	59	48	69	66
PNC	49	54	55	55	51	51	43	29	35	35	30	29	29	23	44	51	51	49	42	42	33	37	57	53	46	64	53	52	54	41	52	54	41
PVTB	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	58	78	78	100	100	100	95	100	100	100	100	100
PBKS	37	40	42	42	30	28	30	27	30	40	40	38	25	38	37	42	48	48	40	34	40	50	70	70	54	70	65	46	57	57	54	71	66
PVN	100	100	100	100	42	42	26	100	100	100	100	68	68	51	100	100	100	100	68	68	51	80	80	40	100	100	41	100	100	50	100	100	53
RF	39	50	62	62	58	50	58	27	49	52	52	49	45	49	36	56	60	60	54	50	54	31	70	69	39	73	71	39	66	65	39	66	65
RIGS	42	51	51	51	44	44	43	6	9	11	11	11	8	11	24	35	36	36	35	34	34	36	61	61	48	72	69	25	36	36	25	36	36
SIVB	52	62	62	55	56	56	36	76	100	100	89	100	100	80	76	100	100	91	100	100	80	37	45	45	55	62	50	65	72	69	65	75	70
SKYF	40	51	62	62	62	44	62	18	32	49	49	49	30	49	29	46	62	62	62	41	62	34	77	77	42	84	83	31	70	70	31	70	70
TSFG	46	58	63	63	63	52	63	13	22	30	30	30	21	30	29	47	56	56	55	44	55	33	67	67	46	83	81	27	52	52	27	52	52
SWBT	33	41	41	41	37	37	36	11	17	21	21	21	17	21	22	36	38	38	37	35	35	34	60	60	41	65	64	27	42	42	27	42	42

STT	44	62	62	50	62	62	37	24	40	40	27	38	38	20	38	56	56	43	53	53	29	28	42	39	39	54	43	38	41	32	38	41	32
SBIB	37	40	42	42	37	34	37	19	21	29	29	28	20	28	37	41	43	43	41	37	41	48	73	73	55	71	70	41	53	53	55	71	70
STI	50	80	100	100	100	80	100	33	58	61	58	61	58	58	48	75	85	85	85	75	85	24	72	72	40	91	91	43	68	68	43	68	68
SUSQ	33	38	42	42	39	30	39	23	30	49	49	49	29	49	33	43	54	54	52	38	52	40	74	74	45	74	74	38	65	65	45	76	76
SNV	36	43	48	48	39	30	39	9	17	21	21	18	14	18	23	32	38	38	30	25	30	45	74	71	46	71	63	33	53	47	33	53	47
TCB	23	25	29	29	25	17	25	13	17	26	26	24	14	24	18	23	31	31	27	17	27	47	80	80	41	64	60	36	55	55	36	55	55
TRST	66	75	75	66	72	72	45	40	49	49	40	49	49	31	66	81	81	66	81	81	45	44	47	47	69	69	54	51	51	43	69	69	54
TRMK	38	44	48	48	42	34	42	17	24	35	35	34	22	34	27	39	48	48	45	34	45	39	70	70	46	74	72	34	55	55	34	55	55
USB	58	65	79	79	77	65	77	67	76	76	73	75	75	68	77	80	84	83	81	75	79	47	78	73	56	85	80	76	81	73	76	81	73
UCBH	77	100	100	100	100	100	100	14	24	32	32	32	24	32	77	100	100	100	100	100	100	31	61	61	62	100	100	27	49	49	62	100	100
UB	41	88	88	72	88	88	70	15	62	62	48	62	62	46	28	78	78	63	78	78	59	21	50	50	34	69	69	21	54	54	21	54	54
UBSI	44	51	56	56	51	39	51	32	42	68	68	67	40	67	44	58	72	72	70	51	70	41	74	74	51	84	82	43	75	75	51	88	87
VLY	45	60	61	61	58	55	58	16	31	36	36	36	30	36	31	52	58	58	55	49	55	31	61	61	45	78	73	28	53	53	28	53	53
WB	59	73	75	71	75	73	67	51	63	63	51	63	63	43	66	74	74	68	74	74	60	64	68	60	64	78	69	65	65	52	65	65	52
WFC	48	50	74	74	74	50	74	88	93	100	100	100	93	100	81	83	94	94	94	78	94	82	100	100	67	95	95	91	100	100	91	100	100
WABC	49	58	58	54	51	51	41	31	40	46	43	45	40	42	49	65	65	57	63	63	48	39	53	53	54	68	60	41	51	51	54	68	60
WTNY	31	43	43	42	43	43	41	11	21	24	24	24	21	24	21	38	40	40	40	38	39	29	57	57	36	63	62	24	43	43	24	43	43
WL	47	58	62	62	57	49	57	14	24	31	31	31	22	31	30	47	55	55	52	41	51	35	66	66	48	81	76	30	52	52	30	52	52
WTFC	54	66	67	67	64	62	64	7	10	15	15	15	10	15	31	47	50	50	49	47	48	35	64	64	54	85	83	26	40	40	26	40	40
ZION	39	59	67	67	66	59	66	14	41	44	44	43	40	43	27	56	60	60	57	53	57	27	65	65	37	76	76	24	60	60	24	60	60

Table 4. Efficiency results under all specifications

Visual examination of Table 4 reveals some important features. Only one bank, Citicorp, (CIH) is efficient under all specifications, highlighting the fact that the selection of inputs and outputs and, therefore, the view of what constitutes efficiency in the financial sector, is a matter of importance. This was one of the conjectures that guided this research. Some banks (PVTB, PVN, CMA, UCBH) are 100% efficient under many specifications. In the same way, some banks achieve low scores under most specifications. Take, for example, PVN, which is 100% efficient under 18 specifications, implying that this is an excellent institution. However, its efficiency drops to 26% under A3. This suggests the presence of some weakness in PVN, a subject that will be further explored below. A counter example is NBY, whose DEA scores tend to be low, but becomes 100% efficient under 3 specifications: C1, C3, C13. This indicates that, although NBY can take action to improve its efficiency, it has some strong points that deserve further attention.

Consider now the case of two institutions that achieve the same DEA score under a given specification. An example would be CYN and IFC. They both are 100% efficient under AB123. But differences appear if other specifications are considered. For example, under AB2 CYN achieves 100% efficiency while the same score for IFC is 37%. Under specification C13 CYN is 52% efficient while IFC is 100% efficient. This indicates that the two institutions follow two different paths to efficiency. What is behind their strategies? Answering such a question was another of the objectives of this research.

In summary, the level of efficiency achieved by a particular financial institution depends on the chosen specification, indicating that specification search is delicate and important. In addition, if two financial institutions achieve the same efficiency score under a given specification they may do so following very different patterns of behavior: there is no single path to efficiency in financial institutions. Exploring what is behind a DEA score is the objective of the next three subsections.

3.2. DEA specification searches using multivariate methods

Although visual inspection of Table 4 is a source of important insights, a more formal analysis of the information it contains will be performed. Table 4 will be treated as a matrix with 85 cases, the banks, and 33 variables, the specifications, and analyzed using multivariate statistical methods. The methodological approach will combine PCA, HCA, and Pro-Fit.

The results of applying PCA to Table 4 are shown in Table 5. Four eigenvalues take values larger than one, accounting for 94% of the total variance. The first principal component accounts for 63%

of the variance. The second principal component is also of importance, as it accounts for a further 14%. The variance accounted for drops to 9% in the case of the third component, and to 7% in the case of the fourth component. Component loadings are given in Table 6. In what follows the discussion will be based on these four components.

Component	*Eigenvalue*	*% of variance*	*Cumulative*
PC1	20.833	63.132	63.132
PC2	4.687	14.203	77.335
PC3	3.097	9.384	86.719
PC4	2.367	7.174	93.893
PC5	0.608	1.843	95.737
PC6	0.555	1.683	97.420
PC7	0.223	0.677	98.097

Table 5. PCA results.

	PC1	*PC2*	*PC3*	*PC4*
ab13	.961	-.109	-.014	-.031
ab123	.949	-.160	.057	-.101
ab12	.917	-.204	.264	-.117
ab23	.916	-.280	-.083	-.126
ab3	.904	-.227	-.298	.007
b13	.903	.084	-.162	-.350
b123	.900	.048	-.103	-.397
b23	.894	-.029	-.188	-.387
b12	.881	.005	.082	-.414
b3	.872	-.006	-.361	-.307
b2	.861	-.124	.036	-.396
bc13	.860	.346	-.167	-.084
ab1	.849	.256	.401	.000
ab2	.844	-.388	.180	-.133
ac13	.824	.061	-.128	.490
a13	.822	-.364	.148	.312
abc13	.821	.375	-.131	.032
b1	.816	.377	.254	-.258
bc3	.810	.167	-.536	-.063
a123	.803	-.433	.213	.236
abc3	.766	.184	-.491	.058
a3	.756	-.486	-.177	.349
a1	.752	.008	.549	.297
a23	.749	-.578	.013	.214
a12	.730	-.425	.446	.211
ac3	.693	-.110	-.477	.492
a2	.672	-.611	.250	.155
bc1	.659	.634	.334	-.054
abc1	.641	.635	.340	.026
ac1	.589	.554	.457	.272
c1	.402	.780	.340	.174
c13	.556	.622	-.281	.389
c3	.477	.453	-.591	.393

Table 6. Component score matrix

Component scores were calculated for each bank. The plot of the first and second component loadings for each bank is shown in Figure 1. The plot of the third and fourth component loadings for each bank is shown in Figure 2.

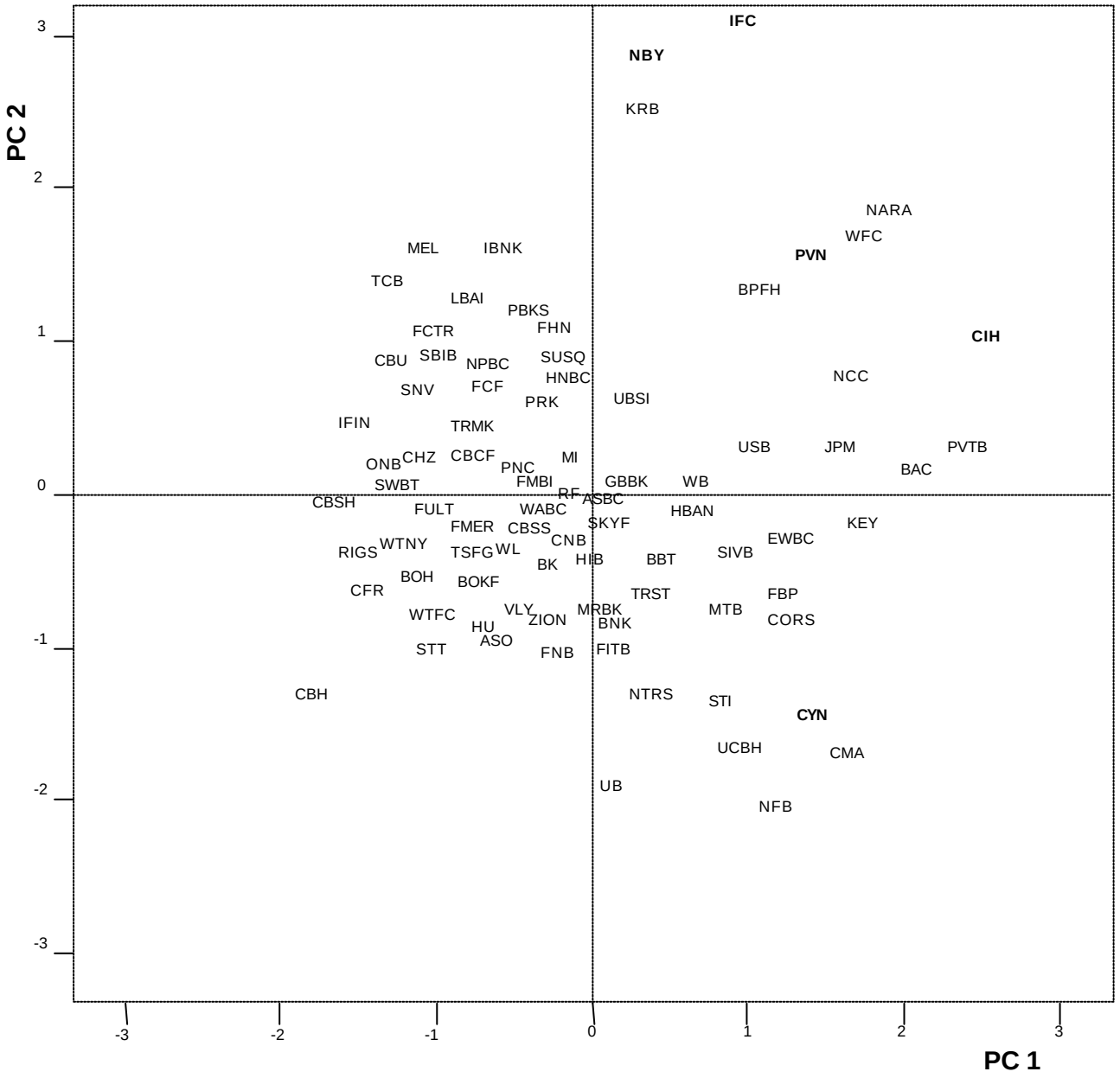

Figure 1. Plot of the first and the second principal component scores. The banks that are discussed in detail in the text have been highlighted in bold.

Figure 2. Plot of the third and the fourth principal component scores. The banks that are discussed in detail in the text have been highlighted in bold.

The banks that achieved full 100% efficiency under a majority of specifications (CIH, PVTB, PVN, CMA, and UCBH) plot towards the right hand side of Figure 1. The banks that consistently under perform plot towards the left hand side of this same figure. It is, therefore, clear that the first principal component can be interpreted as a "global efficiency score". An efficiency ranking of banks can be obtained by simply looking at the ordering on the first component. Usually, efficiency rankings are based on the concept of super-efficiency introduced by Andersen and Petersen (1993), although other ranking methods have also been proposed; Doyle and Green (1994), Sinuany-Stern and Friedman (1998), and Raveh (2000). The advantage of the ranking procedure applied here is that it embeds results from many different specifications, while the alternatives produce a ranking

for each specification. Furthermore, this method permits a ranking of all the banks, whether efficient or inefficient, while under the alternative methodologies only efficient banks can be ranked.

Concentrating now on the second component, the North-South direction in Figure 1, it can be observed that IFC plots towards the top of the figure, while CYN plots towards the bottom. Both are 100% efficient under many specifications. In which way they are different, and what accounts for their achieving full efficiency, will be revealed by attaching meaning to the second principal component. In the same way, interpretation of the position of banks in Figure 2 requires that meaning be attached to the third and the fourth principal components.

A standard way of attaching meaning to principal components is to analyze component loadings. These are given in Table 6. It can be seen there that all loadings associated with the first component are positive, supporting the view that this component gives an overall measure of efficiency. The specifications that achieve the highest first component loadings are AB13, AB123, and AB12. If a combination of inputs and outputs were to be selected in order to produce a global assessment of efficiency, any of these three models would be appropriate.

Specifications that include deposits as an input (C) are salient in the second component, in the sense that they achieve high positive component loadings. The third component appears to be associated with output 3 –loans- versus outputs 2 and 1 – deposits and income-, and the fourth one with input B –physical assets- versus inputs A and C –employees and deposits-.

These results can be visualised by means of Pro-Fit and Cluster analysis. This will be done in the next subsection.

3.3. Results visualization and strategic pattern identification

Each specification generates a DEA score for each bank, and each bank is located in Figures 1 and 2 by means of its component scores. It has just been observed that efficient banks plot towards the right hand side of Figure 1 while inefficient banks plot towards the left hand side of the same figure. This appears to be the case for most specifications. Thus, there appears to be a link between the position of a bank in the figure, as given by the component scores, and efficiency. The relationship between DEA efficiency and component scores can be assessed by means of regression analysis and visualized. For each specification, a regression was run in which the dependent variable was the efficiency value, and the independent variables were the four component scores. Each institution was treated as a case in the regression. In total, 33 regressions were performed. This procedure is

known as Property Fitting (Pro-Fit) analysis; see Schiffman et al (1981). For a given specification, Pro-Fit produces a directional vector on Figures 1 and 2 in such a way that DEA efficiencies grow in the direction of the vector. Directional vectors were calculated for each one of the 33 specifications. Being regression-based, the quality of the representation can be assessed by means of the coefficient of determination, R^2, and the F statistic. These are shown in Table 7. It is to be noticed that values of R^2 are very high, all of them above 0.8, indicating that there is a strong linear relationship between DEA scores and the position of the bank in Figures 1 and 2. The directional vectors are located in Figure 1 and 2 by means of their standardized directional cosines, γ. The standardization transforms the vectors into unit length, using the formula:

$$g_i = \frac{b_i}{\sqrt{\sum_{i=1}^{n} b_i^{\,2}}}$$

The value of the standardized directional cosines, -γ_1, γ_2, γ_3, and γ_4- and of their level of significance, are also shown in Table 7. Pro-Fit vectors have been superimposed on component plots in Figures 3 and 4.

Model	Directional cosines				F	Adj R2
	g_1	g_2	g_3	g_4		
a1	13.05 (32.016)**	0.15 (0.359)	9.52 (23.359)**	5.15 (12.643)**	432.66	0.954
a12	13.61 (31.366)**	-7.92 (-18.263)**	8.31 (19.148)**	3.94 (9.077)**	441.61	0.955
a123	16.05 (27.821)**	-8.65 (-15.003)**	4.26 (7.382)**	4.72 (8.188)**	280.16	0.930
a13	16.69 (27.266)**	-7.38 (-12.056)**	3.00 (4.893)**	6.34 (10.361)**	255.02	0.924
a23	16.47 (27.759)**	-12.71 (-21.411)**	0.28 (0.474)	4.70 (7.920)**	322.98	0.939
a2	14.09 (20.249)**	-12.82 (-18.428)**	5.25 (7.544)**	3.24 (4.662)**	207.06	0.908
a3	17.56 (33.784)**	-11.29 (-21.710)**	-4.11 (-7.907)**	8.11 (15.591)**	479.58	0.958
b1	19.41 (29.614)**	8.96 (13.667)**	6.04 (9.211)**	-6.14 (-9.368)**	309.09	0.936
b12	22.01 (36.650)**	0.13 (0.217)	2.05 (3.409)*	-10.36 (-17.253)**	413.14	0.952
b123	22.89 (56.302)**	1.22 (2.994)*	-2.62 (-6.451)**	-10.09 (-24.817)**	959.08	0.979
b13	22.93 (47.663)**	2.14 (4.458)**	-4.12 (-8.567)**	-8.90 (-18.499)**	676.81	0.970
b23	22.67 (66.740)**	-0.74 (-2.167)	-4.77 (-14.050)**	-9.82 (-28.893)**	1372.77	0.985
b2	21.09 (26.489)**	-3.04 (-3.823)*	0.89 (1.114)	-9.71 (-12.194)**	216.55	0.911
b3	22.03 (63.386)**	-0.16 (-0.472)	-9.12 (-26.238)**	-7.77 (-22.356)**	1301.55	0.984
ab1	18.78 (32.861)**	5.68 (9.932)**	8.87 (15.527)**	-0.01 (-0.015)	354.90	0.944
ab12	19.23 (43.951)**	-4.28 (-9.780)**	5.54 (12.664)**	-2.45 (-5.610)**	554.80	0.963
ab123	0.98 (34.715)**	-0.16 (-5.848)**	0.06 (2.093)	-0.10 (-3.708)*	314.36	0.937
ab13	0.99 (34.118)**	-0.11 (-3.856)*	-0.01 (-0.485)	-0.03 (-1.115)	295.09	0.933
ab23	0.94 (33.492)**	-0.29 (-10.219)**	-0.09 (-3.049)*	-0.13 (-4.601)**	314.15	0.937
ab2	0.88 (25.723)**	-0.41 (-11.830)**	0.19 (5.477)**	-0.14 (-4.061)**	212.03	0.909
ab3	0.92 (38.851)**	-0.23 (-9.740)**	-0.31 (-12.828)**	0.01 (0.308)	442.24	0.955
c1	0.42 (12.381)**	0.82 (24.042)**	0.36 (10.483)**	0.18 (5.371)**	217.51	0.912
c13	0.58 (18.293)**	0.65 (20.443)**	-0.29 (-9.237)**	0.40 (12.796)**	250.41	0.922
c3	0.49 (16.926)**	0.47 (16.054)**	-0.61 (-20.950)**	0.41 (13.918)**	294.21	0.933
ac1	0.61 (21.030)**	0.57 (19.770)**	0.47 (16.320)**	0.28 (9.717)**	298.46	0.934
ac13	0.85 (29.815)**	0.06 (2.204)	-0.13 (-4.631)**	0.51 (17.730)**	307.40	0.936
ac3	0.71 (32.019)**	-0.11 (-5.084)**	-0.49 (-22.043)**	0.50 (22.729)**	513.40	0.961
bc1	0.68 (26.533)**	0.65 (25.533)**	0.34 (13.450)**	-0.06 (-2.170)	385.39	0.948
bc13	0.91 (23.751)**	0.37 (9.557)**	-0.18 (-4.614)**	-0.09 (-2.311)	170.52	0.890
bc3	0.82 (46.866)**	0.17 (9.667)**	-0.54 (-31.009)**	-0.06 (-3.648)*	816.18	0.975
abc1	0.66 (21.869)**	0.66 (21.656)**	0.35 (11.608)**	0.03 (0.890)	270.70	0.928
abc13	0.90 (18.010)**	0.41 (8.221)**	-0.14 (-2.874)*	0.03 (0.692)	100.17	0.825
abc3	0.82 (18.690)**	0.20 (4.487)**	-0.53 (-11.982)**	0.06 (1.423)	128.76	0.859

** Significant at the 0.01 level. * Significant at the 0.05 level

Table 7. Pro-Fit Analysis. Linear regression results

Table 7 about here

Figure 3. Pro-Fit lines and cluster results in the first two component plot.

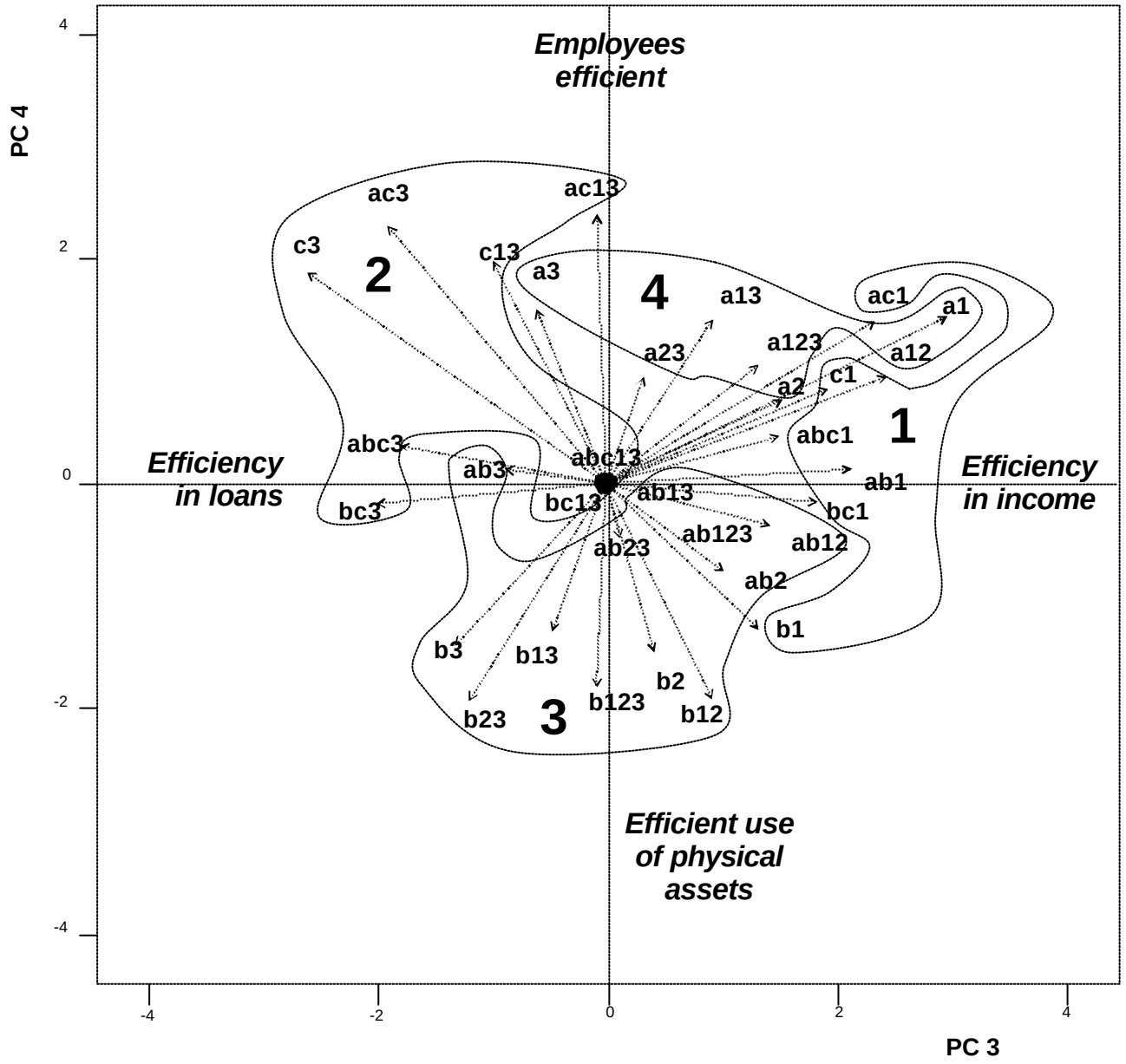

Figure 4. Pro-Fit lines and cluster results on the third and the fourth components plot.

If efficiencies produced by two different specifications are highly correlated, their associated Pro-Fit vectors will plot next to each other. In the same way, if the efficiencies generated by particular specifications are highly correlated with a particular principal component score, the Profit vector will plot in the direction of the axis associated with the given component. The length of the projection of the Pro-Fit vector reflects its relevance in the interpretation of the particular figure. The longer the vector, the more agreement there is between the ordering of the banks in the representation and the efficiency values obtained from the specification.

Pro-Fit vectors form a fan in Figure 3. All vectors point in the direction in which efficiency grows. Each one of the 33 vectors indicates the way in which efficiency grows under a particular

specification. Most vectors point in the direction of the first principal component. This confirms the observation that the first principal component gives an "overall measure of the efficiency" of a bank, and that an ordering along the first principal component produces an efficiency ranking of institutions. Since the first principal component accounts for 63% of the variance, we conclude that the ranking of banks along this component, which is a ranking of overall efficiency, is the most salient characteristic of the data.

A set of vectors, containing deposits as an input, is clearly associated with the second principal component, as they all point towards the top of Figure 3. Deposits are treated as inputs under the production modeling philosophy. We conclude that the second principal component distinguishes between the two basic approaches to banking efficiency: intermediation and production. The second principal component, it has to remembered, accounts for 14% of the variance.

Using similar considerations, we can associate the value of the third principal component – that accounts for 9% of the variance- with the decision to use specifications containing as a sole output the value of loans (3), since a fan that includes only 3 as an output can clearly be discerned on the left hand side of Figure 4. On the right hand side of this figure we find vectors associated with deposits as an output (2) and with income (1). This suggests that the third principal component distinguishes between two banking strategies: collecting deposits, or making loans.

Finally, in Figure 4 it can be seen that the fourth principal component discriminates between specifications that include physical assets as an input (input B) and those contain the number of employees (input A). It is clear that vectors that contain input B in their definition point towards the bottom of Figure 4, while those that contain input A (employees) in their definition point towards the top of the figure. We, therefore, interpret the fourth component as asset utilization efficiency, distinguishing between an orientation towards efficient use of human resources and efficient use of physical assets. This component explains 7% of the variance.

Summarising, in order to describe how a bank achieves a given level of efficiency, we must take into account four independent aspects: first, and most importantly, an overall measure of efficiency; second, we make a distinction between those banks whose efficiency is best seen in the light of the production model, and the banks whose efficiency is best seen under the intermediation model; third, we must take into account whether the bank follows a strategy directed to making loans or to collect deposits; finally, there some banks specialise in the efficient use of human resources while others specialise in the efficient use of physical assets. There may be other features that help to

discriminate between the various banks in terms of efficiency, but they have not been explored, as these four characteristics account for 93% of the variation in efficiency.

All the above discussion has been based on the interpretation of two dimensional projections of a four dimensional data set. Each Pro-Fit vector is plotted in a four dimensional space, and it would be appropriate to assess if the groups that are observed on the projections are true reflections of the groups that exist in the space. For this reason Pro-Fit analysis has been supplemented with Hierarchical Cluster Analysis (HCA). If equivalent specifications exists, they will group into clusters, and if specifications within a cluster share something in common, the analysis will reveal it, with the added bonus that model simplification will naturally follow.

Efficiencies in Table 4 have been taken as inputs for HCA and clustered using Ward's method with Euclidean distances. This method maximizes within group homogeneity and between group heterogeneity. Since we are interested in finding out up to what point two different specifications are equivalent, we clustered variables (specifications) and not cases (banks). The dendrogram can be seen in Figure 5.

Specifications group neatly into four clusters in Figure 5. These clusters have been superimposed in Figures 3 and 4, and have been labeled 1, 2, 3, and 4. Clusters 1 and 2 group together at a higher level of clustering. Clusters 3 and 4 also group together at a higher level of clustering.

Clusters 1 and 2 are located at the North of Figure 3, grouping specifications whose Pro-Fit vectors point up and to the right of the figure. With the exception of specifications B1 and AB1, the 14 specifications that form clusters 1 y 2 contain deposits as input (C). Deposits as an input are a standard feature of intermediation models. Cluster 3 y 4 are located on the lower part of Figure 3, and none of the specifications includes deposits (C) as an input. This is consistent with production models of efficiency.

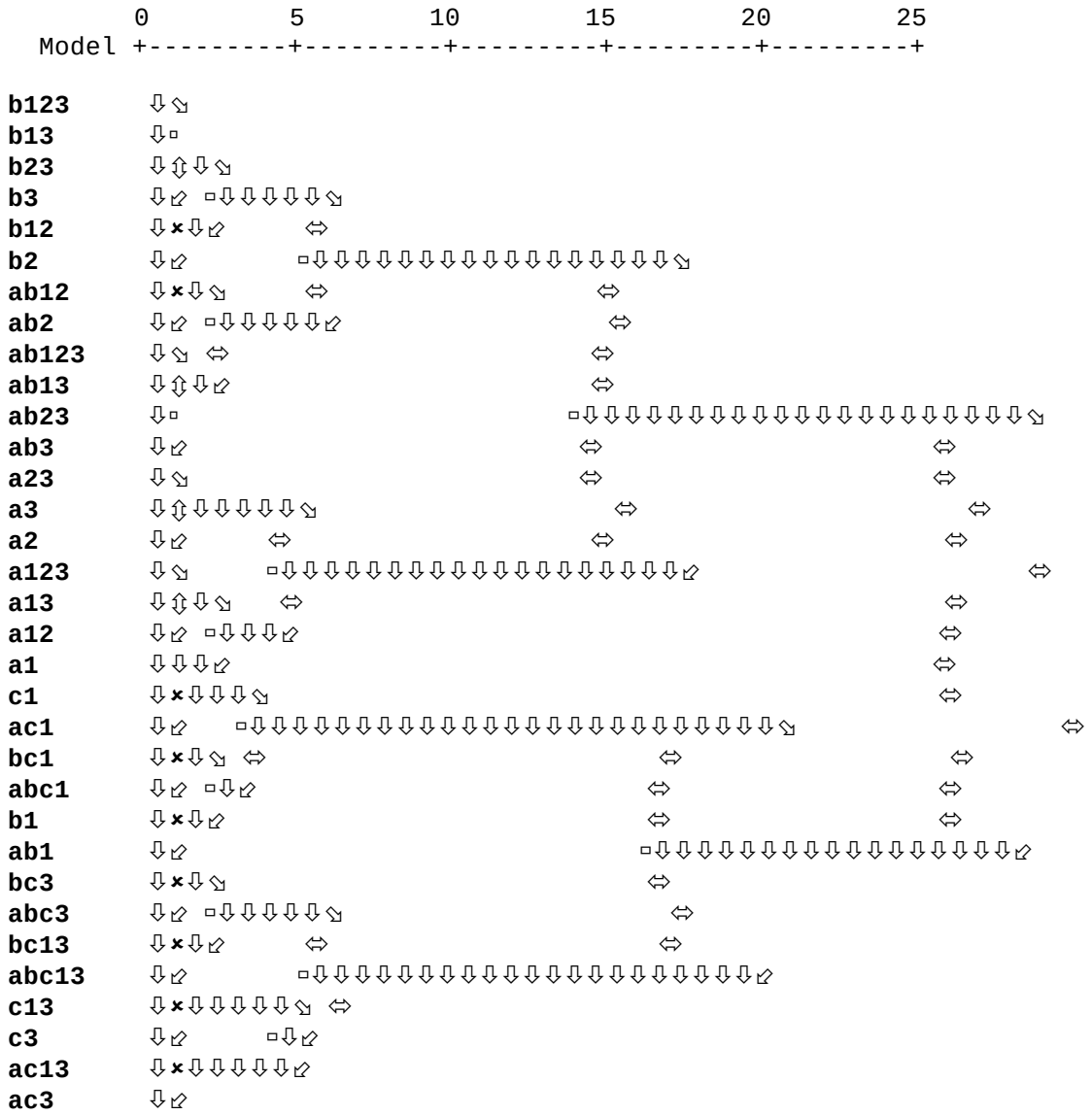

Figure 5. Ward's method. Dendrogram.

The differences that exist between cluster 1 and cluster 2, on the one hand, and cluster 3 and cluster 4, on the other, are made apparent in Figure 4. All the specifications contained in cluster 1 are located at the right hand side of the figure (positive end of the third principal component) and include income (1) as an output. Specifications associated with cluster 2 are located at the left of Figure 4 and include loans (3) as an output. We conclude that cluster 1 contains specifications that are consistent with an intermediation model oriented to income, and that cluster 2 contains specifications that are consistent with an intermediation model oriented towards making loans.

The differences between Clusters 3 and 4 are also apparent in Figure 4. Both clusters are associated with production type models. Cluster 3 is located at the bottom of Figure 4, on the negative side of the fouth principal component. All the specifications contained in it contain input B, physical assets. Cluster 4 is located at the top of figure 4. All the specifications contain number of employees (A) as an input. We conclude that Cluster 3 contains specifications consistent with production models oriented towards efficiency use of physical assets, and that cluster 4 groups specifications consistent with a production model oriented to efficient use of human resources.

It can be argued that specifications contained in a given cluster are largely equivalent in the sense that they produce similar efficiency scores for the various banks. This can guide input and output selection. Each cluster can be represented by a single specification, reducing the total number of possible specifications from 33 to 4. The selected specification could be the most parsimonious one or the most central one within the cluster.

The superimposition of HCA and Pro-Fit results on the component score map clearly reveals the differences between the various modeling approaches. The decision to opt for an intermediation model or for a production model, which is related to the way in which deposits are treated in the specification, will impact on the efficiencies obtained for individual banks. We have also seen that within a particular modeling philosophy, we can have differences based on the orientation towards particular inputs or outputs. In total, we have found 4 alternative definitions of efficiency. This leads to the conclusion that if we want to study the efficiency of a bank, we should not proceed by choosing only one model and only one specification, as this may miss important features of its operations.

3.4. Disecting the efficiency score

It has been argued that there is no single definition of efficiency in the context of banks. Different views of the way in which banks operate, as reflected in the different modeling philosophies will produce different efficiency scores. The combination of PCA, Pro-Fit, and HCA sheds light into the reasons why a particular bank achieves a certain efficiency level. This subject will be further examined in what follows.

Take CYN and IFC, two previously discussed institutions. They both achieve 100% efficiency under 4 specifications: B13, B123, AB123 and AB13. They both appear on the extreme right hand

side of the first principal component in Figure 1. They would both come at the top of an efficiency ranking based on the first principal component. We could just conclude that they are excellent institutions and leave it at that. But it is also to be noticed that under specifications C13, C3, BC13, BC3, ABC13, and ABC3 IFC is 100% efficient but not CYN. All these contain deposits as an input, and are specifications that would be developed under the intermediation modeling philosophy. The specifications that make CYN 100% efficient but not IFC are B12, B23, B2, B3, AB12, AB23, AB2, and AB3. All these specifications contain physical assets (B) in their definition, or employees (A) which leads to the conclusion that CYN owes its position in the league table to the efficient use of its physical and human assets, and that CYN is a good institution from the intermediation point of view.

This discussion can be extended to the differences and similarities of IFC and CYN under the third and fourth principal components. We see that their scores under principal components 3 and 4 are very similar, and this suggests that their only difference is in Principal Component 2, and this has just been discussed.

Systematic analysis of Figures 1 and 2, together with the interpretations provided with the help of Figures 3 and 4 makes it possible to assess the global efficiency of an institution and the strategies under which such global efficiency was achieved. Strengths and weaknesses become apparent. Take, for example, a previously mentioned case: NBY. In Figure 1 NBY plots towards the center of the first component, indicating that its global efficiency is mediocre. Indeed, it only achieves 46% efficiency under specification AB123. It is located at the top of the second principal component, which is consistent with being 100% efficient under specifications C1, C3, and C13, all of them with deposits as an input, and implying that NBY would be only identified as efficient under an intermediation approach. NBY is located very near IFC. It has just been argued that IFC also appeared in a good light under an intermediation approach. The differences between NBY and IFC appear when we examine Figures 2 and 4. In Figure 2, IFC is located towards the most negative side of the fourth principal component, while NBY is located towards the top of the figure, at the positive end of the fourth principal component. In line with the interpretation of the fourth principal component, IFC appears as efficient under specifications that include physical assets (B) as an input and not very efficient when the specification contains the number of employees (A). This is confirmed by the observation that IFC achieves an efficiency of only 12% under specification A2. NBY, on the other hand, appears at its best under specifications that contain input A, the number of employees.

Finally, PVN, another previously discussed bank, appears on the right hand side of Figure 1, implying that it is efficient from the global point of view. Its location in this figure is consistent with intermediation efficiency. In Figure 2, PVN is located towards the extreme right hand side. We notice that in Figure 4, vectors associated with specifications that contain loans (output 3) point on the whole towards the left hand side. This implies that PVN under performs in specifications that contain loans as an output, something that is coherent with the results shown in Table 4. If we were to advise this bank we would recommend more efficiency in the granting of loans.

4. Conclusions

There has been much interest and debate on how to model DEA efficiency in financial institutions. This has extended over the type of model (intermediation or production) that is appropriate, as well as to the selection of inputs and outputs once a modeling philosophy has been selected. We have suggested a specification search strategy that highlights the extent to which two different DEA specifications produce similar results and the reasons why this happens.

The methodology proposed relies on estimating a variety of input/output mixtures and analyzing the results by means of multivariate statistical methods. Particular emphasis is given to data visualization, which is achieved by combining Principal Components Analysis, Property Fitting, and Hierarchical Cluster Analysis.

This approach has been applied to the particular case of American banks. A Principal Component Analysis has made it possible to identify a ranking of banks in terms of global efficiency, which is nothing else than a ranking along the first principal component. Furthermore, we have been able to identify four different views of what constitutes efficiency in a bank. The treatment of deposits as either inputs or outputs- a feature that distinguishes intermediation models from production models- has proven to be a key feature in the modeling of financial institutions, and this information has been captured by the second principal component. Another relevant aspect in the assessment of the differences in banking efficiency is the emphasis on inputs (physical assets versus employees), which is captured by the third principal component. The fourth principal component highlights the institution's orientation towards outputs and separates those institutions that are efficient at granting loans from those that are efficient at taking deposits.

The standard procedure in the assessment of banking efficiency, which starts with an a priori view of what inputs and outputs should be included in the calculation of efficiency should be revised, as different models and specifications can produce different efficiency results for a given institution. A more realistic view would be to accept that efficiency is a multidimensional concept, and that several models ought to be estimated and combined before managerial action is taken to improve the way in which a financial institution works.

Framing DEA results in a multivariate statistical context has allowed us to go behind efficiency as a mere score. It has been possible to offer a global view of the efficiency of an institution which encompasses many specifications; it has made it possible to assess why a particular institution has achieved a given level of efficiency under a given choice of inputs and outputs; and has made it possible to identify the various paths to efficiency followed by different institutions which would, under most studies, have been classified as equivalent but that differ in important aspects of their operations.

Further advantages of the method proposed here is that it creates a natural ranking of institutions in terms of efficiency, and that it highlights the strengths and weaknesses of each institution.

References

Adler, N., Golany, B., 2001. Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. European Journal of Operational Research 132 (2), 18-31.

Aly, H., Grabowsky, R., Pasurka, C., Rangan, N., 1990. Technical, scale, and allocative efficiencies in U.S. banking: an empirical investigation, Review of Economics and Statistics 72, 211-219.

Andersen P., Petersen N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. Management Science 39, 1261-1264.

Athanassopoulos, A.D.,1997. Service quality and operating efficiency synergies for management control in the provision of financial services: Evidence from Greek bank branches, European Journal of Operational Research 98 (2), 300-313

Barr, R.F., Seiford, L., Siems, T., 1993. Envelopment analysis approach to measuring managerial quality of banks, Annals of Operations Research 45, 1-19.

Barr, R. S., Kory, A. K., Siems, T., Zimmel, S., 2002. Evaluating the Productive Efficiency and Performance of U.S. Commercial Banks. Managerial Finance 28 (8), 3-25.

Berg, S.A., Forsund, F.R., Jansen, E.S., 1991. Bank output measurement and the construction of best practice frontiers. Journal of Productivity Analysis 2, 127-142.

Berg, S.A., Forsund, F.R., Jansen, E.S,. 1992. Technical efficiency of Norwegian banks: The non-parametric approach to efficiency measurement. Journal of Productivity Analysis 2 (2), 127-142.

Berg, S. A., Forsund, F.R., Hjalmarsson, L., Suominen, M., 1993. Banking efficiency in the Nordic countries. Journal of Banking & Finance 17 (2-3), 371-388.

Berger, A.N., Humphrey, D.B., 1991. The dominance of inefficiencies over scale and product mix economies in banking. Journal of Monetary Economics 28, 117-148.

Berger, A.N., Humphrey, D.B., 1997. Efficiency of financial institutions: International survey and directions for future research. European Journal of Operational Research 98 (2), 175-212.

Berger, A.N., Mester, L.J., 1997. Inside the black box: what explains differences in the efficiencies of financial institutions? Journal of Banking and Finance 21, 895-947.

Brockett, P.L., Charnes, A., Cooper, W.W., Huang, Z.M. and Sun, D.B., 1997. Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. European Journal of Operational Research 98 (2), 250-268.

Casu, B., Girardone, C., 2004. An Analysis of the Relevance of OBS Items in Explaining Productivity Change in European Banking. EFMA 2004 Basel Meetings Paper.

Charnes, A., Cooper W.W., Huang, Z.M., Sun, D.B, 1990. Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks. Journal of Econometrics 46, 73-91.

Dekker, D., Post, T., 2001. A quasi-concave DEA model with an application for bank branch performance evaluation, European Journal of Operational Research 132, 296-311.

Doyle, J., Green, R., 1994. Efficiency and cross-efficiency in DEA. derivations, meanings and uses. Journal of the Operational Research Society 45, 567-578.

Elyasiani, E., Mehdian, S.M., 1992. Productive Efficiency Performance of Minority and Nonminority-Owned Banks. A Nonparametric Approach. Journal of Banking and Finance 16, 933-48.

Elyasiani, E., Mehdian, S.M., 1995. The Comparative Efficiency Performance of Small and Large U.S. Commercial Banks in the Pre- and Post-Deregulation Eras. Applied Economics 27 (November), 1069-79.

Emrouznejad, A. and Thanassoulis, E. 1996. An extensive bibliography of Data Envelopment Analysis. Journal Papers. Working paper. Warwick Business School.

English, M., Grosskopf, S., Hayes, K., Yaisawarng, S., 1993. Output allocative and technical efficiency of banks. Journal of Banking and Finance 17 (April), 349-66.

Ferrier, G., Lovell, C.A.K., 1990. Measuring Cost Efficiency in Banking. Econometric and Linear Programming Evidence. Journal of Econometrics 46, 229-45.

Hartman, TE., Storbeck, JE., Byrnes, P., 2001. Allocative efficiency in branch banking. European Journal of Operational Research 134, 232-242.

Kumbhakar, S.C, Lozano Vivas, A., Knox Lovell, C., Hasan, I., 2001. The Effects of Deregulation on the Performance of Financial Institutions. The Case of Spanish Savings Banks. Journal of Money, Credit and Banking 33 (1), 101-120.

Kuosmanen, T., Post, T., 2001. Measuring economic efficiency with incomplete price information. with an application to European commercial banks. European Journal of Operational Research 134, 43-58.

Lovell, C.A.K., Pastor, J.T., 1997. Target setting. an application to a bank branch network. European Journal of Operational Research 98, 290-299.

Lozano, A., 1998. Profit efficiency for Spanish savings banks, European Journal of Operational Research 98 (2), 381-394.

Luo X., 2003. Evaluating the profitability and marketability efficiency of large banks. An application of data envelopment analysis. Journal of Business Research 56, 627-635.

Mester, L.J., 1989. Testing for expense preference behaviour. mutual versus stock savings and loans. The Rand Journal of Economics 4, 483-498.

Mester, L.J., 1997. Measuring efficiency at U.S. banks. Accounting for heterogeneity is important. European Journal of Operational Research 98(2). 230-242.

Miller, S. M., Noulas A.G., 1996. The Technical Efficiency of Large Bank Production. Journal of Banking and Finance 20 (April), 495-509.

Ondrich, J., Ruggiero, J., 2001. Efficiency measurement in the stochastic frontier model. European Journal of Operational Research 129, 434-442.

Oral, M., Yolalan, R., 1990. An Empirical Study on Measuring Operating Efficiency and Profitability of Bank Branches. European Journal of Operational Research 46, 282-294.

Pastor, J.M., Perez, F., Quesada, J., 1997. Efficiency analysis in banking firms. An international comparison. European Journal of Operational Research 98, 395-407.

Pastor, J., Ruiz Gomez, J.L., Sirvent, I., 2002. A statistical test for nested radial DEA models. Operations Research 50 (4), 728-735.

Pedraja Chaparro, F., Salinas Jimenez, J., Smith, P., 1999. On the quality of the Data Envelopment Analysis model. Journal of the Operational Research Society 50, 636-645.

Premachandra, I.M., 2001. A note on DEA vs principal component analysis. An improvement to Joe Zhu's approach. European Journal of Operational Research 132, 553-560.

Raveh, A., 2000. The Greek banking system. reanalysis of performance. European Journal of Operational Research 120, 525-534.

Saha, A., Ravisankar, T.S., 2000. Rating of Indian commercial banks. A DEA approach. European Journal of Operational Research 124, 187-203.

Schaffnit, C., Rosen, D., Paradi, J.C., 1997. Best practice analysis of bank branches. An application of DEA in a large Canadian bank. European Journal of Operational Research 98 (2), 270-290.

Schiffman, J.F., Reynolds, M.L. and Young, F.W., 1981. Introduction to Multidimensional Scaling. Theory, Methods and Applications. Academic Press, London.

Sealey, C.W., Lindley J.T., 1977. Inputs, outputs, and theory of production cost at depository financial institutions. Journal of Finance 32, 1251-1266.

Seiford, L.M., 1996. Data Envelopment Analysis: the evolution of State of Art (1978-1995), Journal of Productivity Analysis, 7, 99-137

Seiford, L.M., Zhu, J., 1999. Profitability and marketability of the top 55 U.S. commercial banks Management Science 45 (9), 1270-1288.

Serrano Cinca, C., Mar Molinero, C., 2004. Selecting DEA specifications and ranking units via PCA. Journal of the Operational Research Society 55, 521-528.

Serrano Cinca, C.; Fuertes Callén, Y., Mar Molinero, C., 2005. Measuring DEA Efficiency in Internet companies. Decision Support Systems 38 (4), 557-573.

Sherman, D., Gold, F., 1985. Branch Operating Efficiency. Evaluation with Data Envelopment Analysis. Journal of Banking and Finance 9, 297-315.

Sinuany-Stern Z., Friedman L, 1998. DEA and the discriminant analysis of ratios for ranking units. European Journal of Operational Research 111, 470-478.

Soteriou, A., Zenios, S.A., 1999. Operations, quality and profitability in the provision of banking services. Management Science 45 (9), 1221-1238.

Thompson, R.G., Brinkmann, E.J., Dharmapala, P.S., Gonzalez-Lima, M.D., Thrall, R.M., 1997. DEA/AR Profit-Ratios and Sensitivity of 100 Largest U.S. Commercial Banks. European Journal of Operational Research 98 (2), 213-229.

Tortosa-Ausina, E., 2002. Exploring efficiency differences over time in the Spanish banking industry. European Journal of Operational Research 139, 643–664.

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

Wheelock D., Wilson P.W., 1999. Technical progress, inefficiency, and productivity change in US banking, 1984– 1993. Journal of Money, Credit, Banking 31 (2), 212-34.

Zhu, J., 1998. Data envelopment analysis vs. principal component analysis. An illustrative study of economic performance of Chinese cities European Journal of Operational Research 111, 50-61