Bank cost efficiency and output specification*

Emili Tortosa-Ausina

Universitat Jaume I

October, 2000

Communications to: Emili Tortosa-Ausina, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain. Tel.: +34 964728606, fax: +34 964728591, e-mail: tortosa@uji.es

*I would like to manifest my gratitude to Fundación Caja Madrid, Fundació Caixa Castelló and Generalitat Valenciana for financial support, and to the Instituto Valenciano de Investigaciones Económicas, for kindly providing data. Manuel Illueca provided some help and comments which contributed to improve this paper and must be imperatively acknowledged.

Abstract

There is a longstanding controversy over precisely what it is that banks produce. However, there is little evidence on the sensitivity of bank cost efficiency results when different output measures are applied. In the case of the Spanish banking system, this topic remains virtually unexplored. This paper does exactly that. In particular, we compare nonparametric efficiency scores yielded by two output measures, one of them mostly identified with the asset approach and the other one considering also deposits as output. Results show that distributions of efficiency scores, estimated nonparametrically by means of kernel smoothing, vary much. In addition, firms' positions relative to the mean change according to either output definition, and results do not remain constant over time.

Keywords: banking, nonparametric density estimation, kernel smoothing, transition probability matrix, X-efficiency, cost efficiency

JEL Classification: C14, C30, C61, G21, L5

1 Introduction

In defining the inputs and outputs of banks for the study of efficiency issues, one must choose among several models of the banking firm developed by the literature. There is no consensus on this point, alike when choosing among different techniques to measure efficiency. The latter, though, is a more technical problem, basically linked to the intrinsic features of the models under consideration, either econometric or deterministic, and it is often the case that those differences are the primary cause of differing results.[1]

However, defining different inputs and, particularly, outputs could sometimes imply a more relevant choice, as this decision involves measuring different aspects of the banking firm. Indeed, it is difficult to fully capture the wide range of activities banks produce, due to their multiproduct nature, and the literature has developed different models to do this which do not always find a consensus. Thus, when assessing efficiency issues results differ not only because of the different technique to estimate efficiency but also, by large, because of our beliefs on what banking companies produce.

This turns out to be a problem of major importance, as it could be the case that, under the views of certain output definition, a firm were labelled as inefficient and, simultaneously, according to a different model of the banking firm, it were classified as efficient or, at least, more efficient. Thus, some firms could be mislabelled as inefficient only because we are measuring diverse aspects of the banking firm, specially if the models differ substantially (as it happens in some cases).

Consequently, our conclusions relative to the efficiency and, probably, the competitive viability of some firms in the industry could be somewhat questionable. At least, our final comments should be followed by "according to our definition of bank output" or so, as our model will hardly capture the whole range of products and services provided by the banking firm.

From the excellent survey by Berger and Humphrey (1997) we may infer that, precisely, comparing different output measures for the financial institutions is a research topic to be studied more profoundly, as the attention given to this problem, compared to other research issues, has been minor. Yet, concerns on what banking companies produce turn out to be of chief importance when considering also the strong shifts undergone recently by many banking systems—specially in the Western European area—such as

[1]These are basically the conclusions of the most famous comparison between stochastic and deterministic techniques, by 1990.

deregulation, technological advances, internationalization, etc. As a result, some banking companies could be choosing different and less regulation-conditioned product mixes.[2] In such circumstances, the output definitions should be capable enough to capture these trends.

In the case of the Spanish banking system, this gap in the literature is paralleled. Most studies concerns are strongly linked to the technique used and its appropriateness which, anyway, deserves all our interest. However, virtually only the study by Grifell-Tatjé et al. (1993) considered how different output specifications could bias results, coming to the overall conclusion that, indeed, efficiency scores were sensitive to such varying specifications. The study was highly interesting, as it provided evidence for different definitions of bank output—employed by 8 different research studies—despite being applied only to savings banks, a group of firms currently accounting for less than 35% of total assets in the industry. In the case of the Spanish banking companies, identifying what exactly banks produce turns out to be an issue of major importance, largely because of the major shifts undergone by the industry.

Other research studies that compare different output definitions (not to the Spanish case) are the ones by Kuussaari and Vesala (1995), Berger and Leusner (1997), Berg et al. (1992), Kuussaari (1993), Favero and Papi (1995) or Hunter and Timme (1995). Some of them perform a comparison between the production and intermediation approaches to output measurement, others compare results for outputs measured by numbers of accounts vs. the financial values in these accounts. Some of them make comparisons awarding different nature (input/output) to the deposits. The comparisons of results, though, is in most cases done in the same way: computing the average R_{RANK} for the several approaches under consideration.

The attempts of our paper are partly similar, as we will try to assess how different output measures bias efficiency scores. However, our approach to do it will be completely different, with the basic attempt to identify how the *entire* distribution of efficiency scores behaves. In particular, the technique employed will enable us to assess whether, according to two different models of exactly what it is that banks produce, efficiency scores change their relative positions according to both distributions, and if such distributions match each other. In other words, knowing how an average behaves (in this

[2] This has been occurring for most Spanish commercial and savings banks (Pérez and Tortosa-Ausina, 2000).

case mean efficiency, relative to either output specification) reveals nothing of extreme parts (high and low) of the distribution of efficiency scores. Correlations help somewhat but, again, they are simply a statistic which misses a large amount of meaningful information. In order to draw more accurate—and tight—conclusions, distributions must be analyzed more carefully.

The analysis will be performed by considering two different models to approach output measurement: the asset approach,[3] and another one treating deposits also as outputs and which is closer to the value-added approach.[4] Not considering the production approach involves ignoring the banking facet of services production. However, the available databases do not provide information relative to this point.

The study proceeds as follows. Section 2 deals with the different methods to measure bank output. Section 3 estimates cost efficiency scores for the Spanish banking industry applying nonparametric techniques at every period and according to different output measures. The resulting distributions are compared in section 4 via nonparametric density estimation. Section 5 assesses whether firms positions relative to the mean vary according to either output measure, by means of nonparametric bivariate density estimation. Finally, section 6 concludes.

2 A long-standing disagreement

There are two chief approaches to the choice of how to measure the flow of services provided by financial institutions. Under the "production" approach the banking entities are primarily deemed as service producers for account holders whereas, under the "intermediation" approach, the financial institutions are thought of as primarily intermediating funds between savers and investors. Consequently, the former considers the output as being made up by the number and type of transactions performed or processed documents during a certain period, whereas only the physical inputs are considered (labour and capital). The rationale for this choice would consist of the banking firm being treated only as a service producer for the depositors. On the other hand, under the views of the intermediation approach bank outputs consist of the money value[5]

[3]See Klein (1971) and Sealey and Lindley (1977).

[4]See Berger et al. (1987), Berger and Humphrey (1992) or Clark (1996).

[5]Kolari and Zardkoohi (1987) debate about the convenience of using money value instead of number of accounts, loans, etc. Yet, some studies show that in the case of the Spanish savings banks the number of accounts and their money value are highly correlated, both for loans and deposits (Grifell-Tatjé and

in earning assets, primarily, although deposits play a role which will be discussed below, whereas inputs would comprise not only labour and capital, but also deposits and the financial costs they involve.[6]

It would be desirable applying a *dual* approach capturing both the service production and intermediary nature of banking companies. However, the production approach requires some information which is not publicly available.This prevents us, and many others, from using it.[7] In addition, such an approach may be somewhat better for evaluating the efficiencies of bank branches, whereas the intermediation approach may be more appropriate for evaluating entire financial institutions. But, there are some items which can be attached to the service production nature of the banking firm, as savings deposits or, to some extent, cash and balances with the central bank (despite of its partly compulsory nature; this point is more arguable).

In particular, the asset, user cost, and value-added methods[8] to define bank output differ, among other aspects, in the role attached to deposits. The asset (Sealey and Lindley, 1977) and user cost (Hancock, 1985, 1991) approaches treat inputs and outputs in a mutually exclusive way. The former contemplates banks only as financial intermediaries between liability holders and those who receive bank funds. The latter classifies the different asset and liability categories as inputs or outputs depending on their net contribution to bank revenue. On the other hand, under the views of the value-added method (Berger et al., 1987) liabilities may have simultaneously input and output characteristics. More precisely, all those categories having substantial value added are employed as important outputs.

The appropriateness of each method varies according to different circumstances, and the user cost method might be, in some cases, difficult to implement. This is the case for our database, as this method requires—among some additional information—data on interest and other income received for the different asset categories, or paid for liability categories, which are unavailable at disaggregate level. Precisely, the two main criticisms to this approach consist both of the difficulties collecting data and the practice

Lovell, 1996).

[6] Moreover, it should be pointed out that, in general, financial costs more than double operating costs (Humphrey, 1992).

[7] As Colwell and Davis (1992) suggest, at the core of the problem might lie both the complexity of banking as an activity and—specially—poor data.

[8] These three distinct approximations to bank output are accurately explained in Berger and Humphrey (1992).

of subsidization, implying low reliability of prices and available revenues.[9] Besides, the value-added and user cost approaches give roughly similar results, at least in some occasions.[10]

Consequently, our first approach to output measurement will be mostly identified with the intermediation and, more closely, the asset approach, as it will treat as bank output only earning assets. The second approach will consider, though, that most banks raise a substantial portion of their funds through produced deposits and provide liquidity, payments, and safekeeping services to depositors to obtain these funds. Accordingly, it will consider a different output definition, basically because of treating savings deposits both as inputs and outputs, alike other studies applying the value-added method.[11]

All variables are described in table 1, which also reports some basic information provided by the Spanish commercial banks association (AEB, Asociación Española de Banca) and the Spanish savings banks association (CECA, Confederación Española de Cajas de Ahorro). The firms which were not in continuous existence over the sample period 1985–97 were dropped, and banks were backward merged in order to have the same number of firms at every year. Although this could seem an important loss of data, our sample always covered around 90% of total industry assets.

3 Nonparametric estimates of cost efficiency according to different output measures

The second (not in importance) source of debate and controversy when studying bank efficiency issues arises from the choice of technique for its measurement. The purpose of this study, though, is not to make a comparison between different techniques and, accordingly, we have chosen only one. In particular, the nonparametric ADEA (Allocative Data Envelopment Analysis) technique[12] to measure cost efficiency has been selected,

[9] See Favero and Papi (1995).

[10] See Berger and Humphrey (1992).

[11] Yet, this method involves estimating which variables yield value-added enough to be treated as outputs. However, we will follow the findings in Berger and Humphrey (1992) and Berger et al. (1987) and, consequently, the outputs according to this approach will be much the same as those used there. This way to proceed is similar to that of Clark (1996), who also follows Berger and Humphrey (1992). Besides, he points out reasons to also treat as output securities, despite absorbing less than 2% of value-added. Such studies identified the major categories of produced deposits (demand, time, savings) and loans (real state, commercial, installment) as important outputs, whereas purchased funds (federal funds purchased, large CDs, foreign deposits, other liabilities for borrowed money) are thought of only inputs.

[12] See Aly et al. (1990).

Table 1: Definition of the relevant variables (1997)

Variable	Variable name	Definition	Mean	Std Dev
Outputs				
Approach 1 to output measurement				
y_1	Loans‡	All forms of loans to customers	375536	698227
y_2	Other earning assests‡	Securities and loans to financial institutions	373427	914773
Approach 2 to output measurement				
y_1	Loans‡	All forms of loans to customers	375536	698227
y_2	Other earning assests‡	Securities, loans to financial institutions, and cash balances	373427	914773
y_3	Savings deposits‡	Savings deposits (includes also time deposits)	376079	689750
Inputs (common to both approaches)				
x_1	Labour‡	Total labour expenses	10780	20423
x_2	Funding‡	Savings deposits, other deposits, and interbank deposits	721068	1525598
x_3	Capital‡	Physical capital	19931	39705
Inputs' prices (common to both approaches)				
ω_1	Price of labour	labour expenses/number of employees	5.122	1.141
ω_2	Price of funds	financial costs/x_2	0.060	0.183
ω_3	Price of physical capital	(amortizations+other non-interest expenses)/x_3	0.499	0.420

‡In millions of 1990 pesetas.

because of its ability to envelope data quite closely, despite its inability to disentangle inefficiency from random error. Parametric methods do this but, in turn, they must impose a functional form on the distribution of inefficiency which, in principle, involves less flexibility.[13] However, no methodology dominates the other.[14]

This technique estimates efficiency by solving the following program:

$$\begin{array}{lll} Min_{x_{js}} & \sum_{j=1}^{n} \omega_{js} x_{js} & \\ s.a. & y_{is} \leq \sum_{s=1}^{S} \lambda_s y_{is}, & i = 1, \ldots, m, \\ & x_{js} \geq \sum_{s=1}^{S} \lambda_s x_{js}, & j = 1, \ldots, n, \\ & \lambda_s \geq 0, & s = 1, \ldots, S, \\ & \sum_{s=1}^{S} \lambda_s = 1 & \end{array} \tag{1}$$

where firm s uses an input vector $x = (x_1, \ldots, x_j, \ldots, x_n) \in \mathbb{R}_+^n$ available at prices $\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}_+^n$ for producing outputs $y = (y_1, \ldots, y_i, \ldots, y_m) \in \mathbb{R}_+^m$.

Computing the individual cost efficiency scores requires solving program (1) for each s firm and year in our sample. The solution will be given by the x_s^* cost minimizing vector, given the price vector ω_s and outputs vector y_s.

Accordingly, the efficiency scores are given by:

$$ES_s = \frac{\omega_s^{'} x_s^*}{\omega_s^{'} x_s} \tag{2}$$

Similarly, the inefficiency estimates will be given by:

$$IS_s = \frac{1}{ES_s} - 1 \tag{3}$$

which reveals the amount to which firms s costs are increased for performing off the efficient frontier made up of those "best-practice" banks.

Results are reported in tables 2 and 3. They show that things do not behave in the same way when different output definitions are considered. More properly, the

[13]The papers by Ferrier and Lovell (1990) or Resti (1997) provide excellent comparisons among the different types of techniques.

[14]Berger and Humphrey (1997) confirm that, out of 130 applications, more than half employed nonparametric techniques, and 60 were parametric. More recently, Bauer et al. (1998) have suggested nonparametric do not meet their consistency conditions and accordingly should not be used. On the contrary, McAllister and McManus (1993), Mitchell and Onvural (1996) and Wheelock and Wilson (2000) test and reject the translog specification of bank cost functions, and suggest semi-nonparametric or nonparametric methods for estimating bank costs.

moments under consideration (mean, standard deviation) differ much. Probably, the most striking feature is a much dissimilar trend in the efficiency evolution. According to the first approach to output measurement, there has been a sharp increase (rising from 62.87% in 1985 to 80.21% in 1995, for the mean values of the whole industry), whereas the pattern is far more stable according to the second approach (the extreme values are 78.69% and 84.84%).

Yet, conclusions diverge upon the type of firm considered, or firm's size. When deposits are treated also as output (approach 2), savings banks' efficiency is always higher than commercial banks', if simple mean values are analyzed. However, the pattern is reversed—at least until 1993—according to the asset approach. Weighted values show also that firms' size is an issue to account for, as they are always higher, or much higher, than their simple counterparts. This trend further contributes to make us appreciate that, indeed, conclusions drawn at industry level might mask important tendencies at firm level.

Standard deviation helps in this, of course. Its value varies much according to the issues considered (type of firm, approach to output measurement), but fails in recording features as important as, for instance, multiple modes. It could be the case that a group of very efficient and very inefficient firms existed simultaneously and were approaching over time. However, the dispersion indicator might remain stable over the analyzed period. In other words, two distributions, with a different amount of multi-modality (one uni-modal, the other one bi-modal, for instance) might have very similar dispersion indicators.

4 Are two moments of the distributions enough?

Nonparametric estimates of density functions overcome such difficulties. In addition, they contribute to shed light on the differences of the distributions according to different output definitions. Some studies comparing efficiency results obtained with outputs measured by numbers of accounts vs. the financial values in these accounts found differences for the mean values (Kuussaari, 1993). Yet, distributions were much the same.[15] However, conclusions were drawn on the frequency distributions of both technical and scale inefficiency scores. These instruments provide a good starting point in order to de-

[15] Although according to Berg et al. (1992), results were much similar also for mean efficiencies.

Table 2: Efficiency evolution, banking firms (1985–97) (*approach 1*)

		1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
	Simple mean	70.46	62.42	67.34	71.38	72.96	72.33	76.92	78.85	78.09	75.12	78.86	68.20	69.04
Commercial banks	Weighted mean	87.19	76.98	84.89	87.75	89.03	88.89	86.15	90.52	81.99	88.92	88.93	78.09	86.54
	Standard deviation	19.80	21.01	21.53	20.76	18.21	19.60	17.27	16.85	17.16	18.47	16.58	19.17	20.12
	Simple mean	54.67	46.21	52.06	56.87	59.65	61.65	71.48	77.28	78.11	76.86	81.68	71.42	68.33
Savings banks	Weighted mean	64.53	60.79	66.60	73.60	76.45	77.36	74.53	85.24	82.06	82.86	85.79	78.45	78.90
	Standard deviation	15.97	17.09	14.91	15.62	14.70	14.35	10.37	10.51	10.49	9.97	10.42	12.58	10.99
	Simple mean	62.87	54.62	59.99	64.41	66.56	67.20	74.30	78.09	78.10	75.95	80.21	69.75	68.70
Total	Weighted mean	79.65	71.15	78.24	82.51	84.22	84.60	81.75	88.51	82.02	86.50	87.68	78.26	83.45
	Standard deviation	19.71	20.86	20.15	19.84	17.90	18.08	14.62	14.18	14.35	15.02	14.03	16.41	16.38
# of commercial banks		54	54	54	54	54	54	54	54	54	54	54	54	54
# of savings banks		50	50	50	50	50	50	50	50	50	50	50	50	50
# of banks (total)		104	104	104	104	104	104	104	104	104	104	104	104	104

Table 3: Efficiency evolution, banking firms (1985–97) (*approach 2*)

		1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
	Simple mean	78.88	75.08	79.11	80.81	81.37	80.98	80.07	81.43	80.67	78.00	81.90	78.10	77.12
Commercial banks	Weighted mean	90.20	87.77	88.51	89.69	90.58	89.14	90.56	91.08	83.07	89.67	88.42	89.40	89.04
	Standard deviation	16.61	14.95	15.84	15.08	13.92	16.18	16.21	15.86	15.72	17.02	14.39	17.22	17.61
	Simple mean	85.57	82.59	84.08	82.17	81.63	83.85	82.29	85.06	84.28	83.53	88.02	88.17	87.76
Savings banks	Weighted mean	86.24	87.03	88.17	88.75	89.03	90.37	85.77	90.37	89.05	89.65	88.95	93.23	93.24
	Standard deviation	9.63	9.55	9.70	10.29	10.39	10.02	11.22	9.80	9.62	9.57	8.08	8.11	8.57
	Simple mean	82.09	78.69	81.50	81.46	81.49	82.36	81.14	83.18	82.40	80.66	84.84	82.94	82.24
Total	Weighted mean	88.95	87.50	88.39	89.34	89.99	89.65	88.88	90.81	85.26	89.66	88.99	90.92	90.73
	Standard deviation	14.11	13.19	13.48	13.01	12.35	13.65	14.08	13.42	13.27	14.21	12.17	14.52	14.98
# of commercial banks		54	54	54	54	54	54	54	54	54	54	54	54	54
# of savings banks		50	50	50	50	50	50	50	50	50	50	50	50	50
# of banks (total)		104	104	104	104	104	104	104	104	104	104	104	104	104

Efficiencies have been estimated for each year separately, and common frontiers for commercial banks and savings banks are specified.

tect, for instance, whether multi-modality exists. However, they are affected by several problems, brilliantly reported by Silverman (1986).[16]

A better way to detect data structure consists of falling back on the kernel method to nonparametrically estimate density functions. This method permits uncovering all features data might hide much accurately than, for instance, an histogram does. Of course, there are other methods to smooth data but this is, by far, the best applicable to most circumstances. The easiness to understand its properties contributes further to make it more popular.[17]

This method consists of—after normalizing[18] efficiency scores—estimating the following density function for each output specification and year (or period):

$$\widehat{f}(x) = \frac{1}{Sh}\sum_{s=1}^{S} K(\frac{x - NES_s}{h}) \tag{4}$$

where S is the number of firms in our sample, NES_s is the normalized efficiency score for firm s, and h is the bandwidth, window width or smoothing parameter, which determines the amount to which data will be smoothed.

K is a kernel function satisfying:

$$\int_{-\infty}^{+\infty} K(t)dt = 1 \tag{5}$$

Kernel's choice consists of several alternatives.[19] In our case we have selected the Gaussian kernel which, in the univariate case we are dealing with is expressed by:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \tag{6}$$

The relevant choice, though, is not the kernel's but, by large, the h's or bandwidth's. While the kernel determines the shape of the bumps when plotting function (4), the smoothing parameter has a different effect, conditioning bumps' width. If h is too

[16] See also Scott (1992), Wand and Jones (1995), Simonoff (1996)), Devroye and Györfi (1985) or Nadaraya (1989).

[17] Along with the histogram and kernel estimator we may find the naive estimator, the nearest neighbour method, the variable kernel method, the orthogonal series estimators, the penalized maximum likelihood estimators, etc.

[18] Or dividing by the mean. Consequently, if the normalized efficiency score of certain firm had a value of 2, it would indicate such a firm is twice efficient than industry average. On the other hand, if such a value were 0.5, it would indicate its efficiency is half of industry average.

[19] Epanechnikov, triangular, Gaussian, rectangular, etc.

small, an excessive number of bumps is generated and data structure is difficult to appreciate; in other words, data are *undersmoothed*. On the other hand, if h is too large *oversmoothing* occurs, and some data features are hidden. What we find under these graphic facts is the traditional trade-off between bias and variance which, indeed, depends on the smoothing parameter: the larger is h, less variance and more bias, and vice versa.

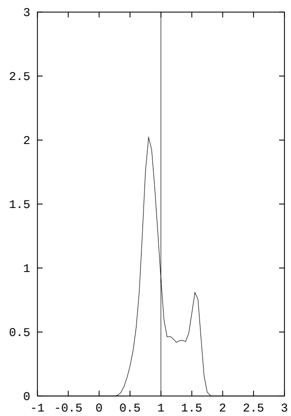
The relevance of this decision has led us to take some cautions on this topic and, finally, to choose the smoothing parameter suggested by Sheather and Jones (1991), based on Park and Marron (1990). It relies on the second generation method *solve-the-equation plug-in*, and its higher performance relative to first generation methods has been verified in further research studies.[20]

Nonparametric density estimates for both output specifications are reported in figures 1 and 2. Splitting them into sub-figures has been done in an attempt to better capture the whole time span of data. The features such figures reveal are manifold but, probably, one of the most interesting ones—for our purposes—lies in that distributions are not, by large, the same. Only in 1991 we find a similar shape and, to some (close) extent, for the period 1991–96, but this is not paralleled for any other period. In addition, we ignore whether changes in firms' relative positions occur among these two specifications. Yet, this question will be properly answered below.

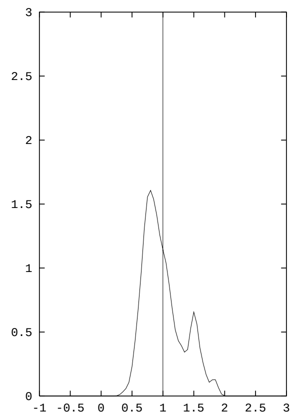
But more information—not revealed by tables 2 and 3—is available. For instance, according to the first approach to output measurement, there has been a sharp fall of dispersion comparing 1985 and 1997. However, what density functions show (figures 1.a and 1.b) is that there was an evident bimodality which has partially diminished. In addition, and this is a trend common to both approaches, it seems that those best-practice firms against which efficiency is assessed are always far enough from all other firms in the sample to form a perceptible mode.

[20]See, for instance, Jones et al. (1996) or the simulation studies by Park and Turlach (1992) or Cao et al. (1994). More details on our bandwidth are available from the papers by Sheather and Jones (1991) and Park and Marron (1990). In addition, Steve Marron's *web*'s page provides the Matlab routine which enables its obtaining (URL: http://www.stat.unc.edu/faculty/marron.html, accessed September, 1999).

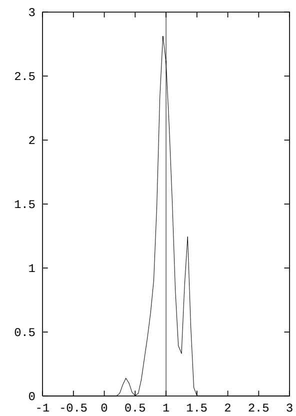
Figure 1: **Normalized efficiency densities (*approach 1*)**



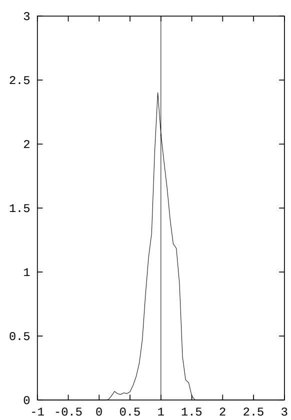
a) 1985



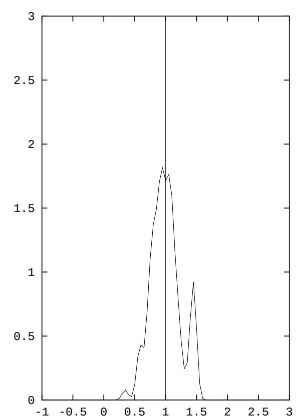
b) 1986–90



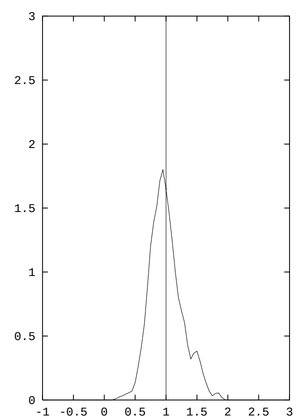
c) 1991



d) 1992–96

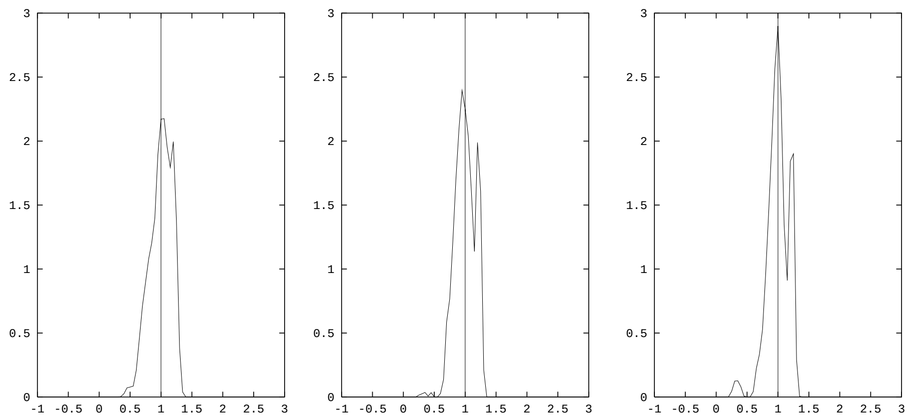


e) 1997

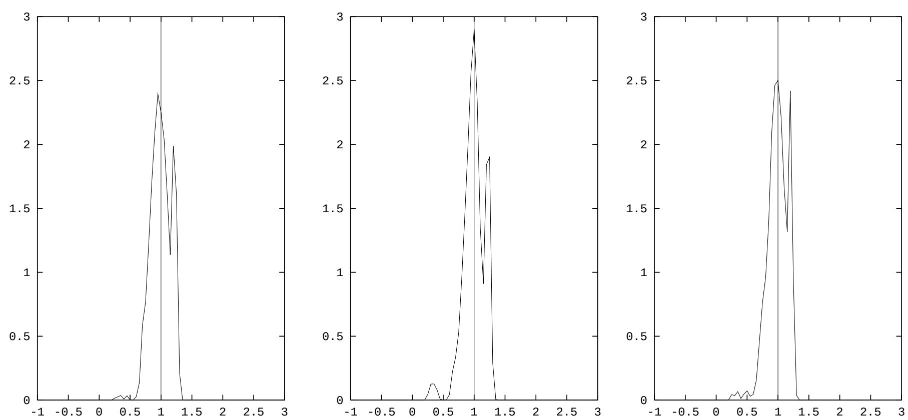


f) 1985–97

Figure 2: Normalized efficiency densities (*approach 2*)



a) 1985 **b) 1986–90** **c) 1991**



d) 1992–96 **e) 1997** **f) 1985–97**

Table 4: Estimated correlations of variables

Period	R_{RANK}
1985	0.559
1986–90	0.677
1991	0.885
1992–96	0.873
1997	0.813
1985–97	0.726

5 Changes in firms' relative positions

That information figures 1 and 2 do not provide relates to firms' relative positions according to either output measures. In other words, we still ignore whether each firm's efficiency score is robust to output specification. Prior research studies approached this issue by computing correlation ranks, and results vary across them, ranging from a R_{RANK}=0.16 (Hunter and Timme, 1995) to R_{RANK}=0.77 (Favero and Papi, 1995).

We add here an additional source of dispersion in results, which is time. It could be the case that efficiency scores were more robust to output specification in some periods and these conclusions varied substantially in others. Some studies[21] have reported evidence on banks' shifts in specialization throughout the 1985–1997 period. Different output measures involve emphasizing different lines of business, and if these do not remain stable over time, we might expect different banks' efficiency scores for either output measures.

Table 4 displays information on estimated correlations between the scores of individual banks with different output definitions for the selected periods. It shows that, indeed, the correlation rank varies sharply along time, as its value for 1985 is 0.559, reaches a peak in 1991 (0.885) and falls partially in 1997 (0.813). Despite the decrease in the last year of the sample, the tendency displays a steady increase over the entire period.

Our main contribution, though, consists of applying a different and more accurate technique to identify changes in firms' relative positions. Drawing conclusions from a single statistic does not give a full view of the facts under consideration. Knowing how this statistic behaves reveals some interesting facts, but there exists a lot of meaningful information it fails to uncover. For instance, although we know that in 1985 there were marked changes in firms' relative positions, we ignore transitions' paths or, in other

[21] See Pérez and Tortosa-Ausina (2000).

words, were firms lied according to either output definition.

The technique to be applied here does exactly that. In particular, transition probability matrices across the output definitions under consideration are estimated, in order to identify firms' mobility. Then, if firms' positions were invariant relative to the mean—as we previously normalized data—these transition probability matrices should be the identity matrix: the distributions are invariant and, in addition, firms' mobility does not exist. On the other hand, if entries off the diagonal were different from 0, then movements across these two distributions would be occurring.

It must be borne that these transition probabilities describe transitions from the second definition of bank output to the first one—mostly identified with the asset approach—. Consequently, we are not quantifying transitions over time. In addition, this type of analysis involves choosing according to some rule the grids, or limits, of states. In this case cells are arrayed in increasing order, with the lower right-hand corner displaying transitions from the most efficient firms to the most efficient firms, according to either output definition. Moreover, the first column displays the number of firms included in each state of relative efficiency, according to the second output definition.

Table 5 displays these transition probabilities from approach 2 to 1, for the periods under study. It shows clearly that the information table 4 provides misses some features. Particularly, the top left-hand entry in table 5.a shows that the less efficient 20% of banking companies according to the second approach to output measurement—with efficiency scores less than 84.7% of the average—remained with efficiencies in that range with probability 0.71 according to the first approach. The remainder 0.29 moved overwhelmingly (0.24) to state 2, including the following 20% of less efficient firms (ranging from 84.7% to 96.9% of average efficiency) and, surprisingly, state 5 of upper relative efficiency (0.05). Yet, it is even more striking noticing how probability abandons almost completely the diagonal in states 2, 3 and 4. Transitions occur always to a lower efficiency state, and a very small amount of probability remains in the diagonal (except in the case of state 4, where all probability moves to other states).

These tendencies are similar for the 1986–90 period. But transitions also occur, although to a far more modest rate, in 1991 and in the 1992–96 period, and more intensively in 1997, despite the high correlation coefficients. It is particularly remarkable how 24% of probability abandons the lower right-hand entry and shifts to the first state of less relative efficiency (table 5.e).

This analysis contributes to shed more light on the issues under consideration, but it has a certain disadvantage, consisting of having to choose, somewhat arbitrarily, the limits of the states of relative efficiency. In other words, we have to discretize a continuous process, and every discretization is arguable.

In order to get rid of such a discretization, we have approached the shifts in firms relative positions from a *continuous* point of view. This would imply taking the number of cells, or states, tending to infinite. Consequently, what we analyze now is the continuous counterpart of the transition probability matrices or, more properly, the *stochastic kernels*.[22] We should think of these stochastic kernels as conditional probability density, as they describe the probability according to one output definition *conditional* on the density according to the other. To word it more precisely, they provide information on transitions across output definitions *conditional* on what we begin with. The estimation of the stochastic kernels is done by nonparametrically estimating bivariate density functions—where each variable is the normalized efficiency score according to either output definition—and then dividing by the implied marginal, as we are attempting to obtain conditional probability. Again, the exercise will be carried out for the selected periods.

Yet, the nonparametric estimation of bivariate density functions has problems similar to those faced by the univariate case. An appropriate way to smooth data is, similarly, kernel smoothing,[23] but the choice of bandwidth turns out to be, in this bivariate situation, particularly difficult. This occurs because in this case the state of the art is in a much preliminary stage. We have selected, as before, the solve-the-equation plug-in approach, based on Wand and Jones (1994),[24] where smoothing parameters (one for each coordinate dimension) are provided which, in general, perform better—because of a much precise balance between bias and variance—than least squares cross validation. However, the advantages of this technique are exactly the same: it enables uncovering all feature data might hide, which any parametrization would omit.

Figure 3 reports results on the three dimensional plots of the stochastic kernels, for the selected periods. Conclusions can be more accurately drawn from contour plots,

[22]See Stokey and Lucas (1989). For applications similar to those we are dealing with, see also Andrés and Lamo (1995), Lamo (2000) or Quah (1996, 1997).

[23]Although in this case the chosen kernel is the Epanechnikov's.

[24]In this case, the computation has been enabled through the S-plus code available from URL: http://lib.stat.cmu.edu/S/kernel, accessed September, 1999.

Table 5: Transition probability matrices across different output definitions

(Number)	Upper endpoint: 0.847	0.969	1.068	1.181	∞
(21)	0.71	0.24	0.00	0.00	0.05
(21)	0.48	0.19	0.05	0.10	0.19
(20)	0.60	0.10	0.15	0.05	0.10
(21)	0.19	0.43	0.10	0.00	0.29
(21)	0.10	0.05	0.05	0.05	0.76

a) 1985

(Number)	Upper endpoint: 0.864	0.952	1.041	1.170	∞
(104)	0.74	0.10	0.06	0.09	0.02
(104)	0.57	0.18	0.08	0.13	0.05
(104)	0.41	0.12	0.14	0.13	0.20
(104)	0.28	0.19	0.11	0.17	0.25
(104)	0.01	0.10	0.11	0.07	0.72

b) 1986–90

(Number)	Upper endpoint: 0.887	0.969	1.039	1.178	∞
(21)	0.76	0.24	0.00	0.00	0.00
(21)	0.14	0.48	0.33	0.05	0.00
(20)	0.05	0.35	0.30	0.30	0.00
(21)	0.05	0.10	0.29	0.52	0.05
(21)	0.00	0.05	0.14	0.05	0.76

c) 1991

(Number)	Upper endpoint: 0.890	0.968	1.049	1.157	∞
(104)	0.73	0.27	0.00	0.00	0.00
(104)	0.30	0.45	0.20	0.03	0.02
(104)	0.09	0.28	0.34	0.29	0.01
(104)	0.03	0.07	0.12	0.46	0.33
(104)	0.03	0.01	0.06	0.16	0.74

d) 1992–96

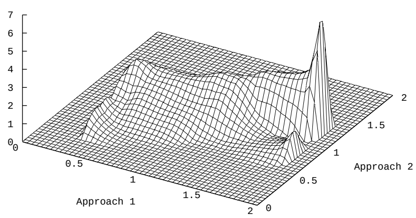
(Number)	Upper endpoint: 0.887	0.969	1.039	1.178	∞
(21)	0.62	0.05	0.10	0.14	0.10
(21)	0.33	0.29	0.10	0.19	0.10
(20)	0.20	0.25	0.05	0.30	0.20
(21)	0.10	0.19	0.29	0.19	0.24
(21)	0.24	0.10	0.10	0.19	0.38

e) 1997

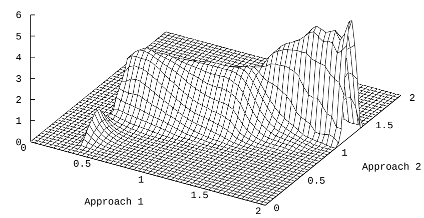
(Number)	Upper endpoint: 0.872	0.965	1.048	1.165	∞
(271)	0.76	0.15	0.03	0.03	0.02
(270)	0.41	0.32	0.16	0.07	0.04
(270)	0.26	0.20	0.25	0.19	0.10
(270)	0.15	0.13	0.13	0.33	0.26
(271)	0.03	0.05	0.08	0.11	0.73

f) 1985–97

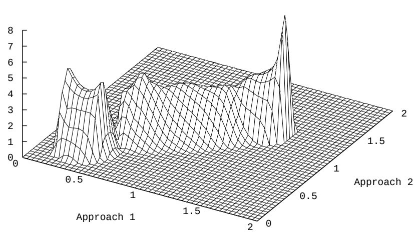
Figure 3: Efficiency transitions across different output definitions



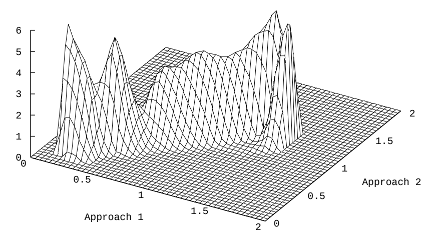
a) 1985



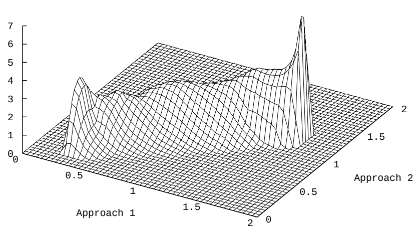
b) 1986–90



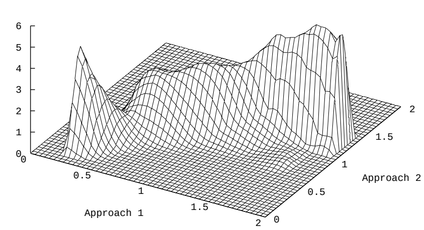
c) 1991



d) 1992–96

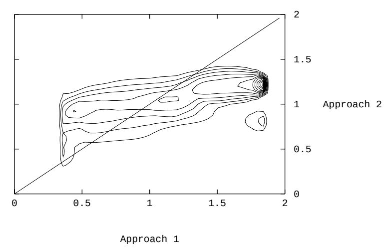


e) 1997

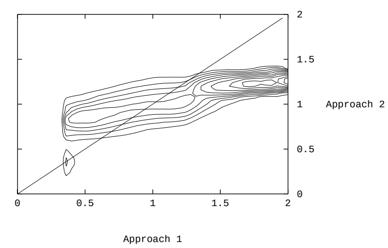


f) 1985–97

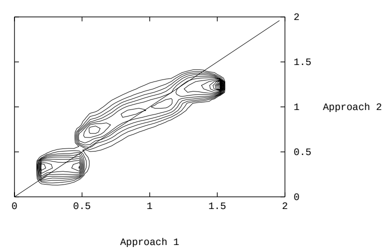
Figure 4: Efficiency transitions across different output definitions (contour plots)



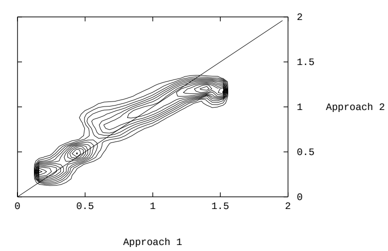
a) 1985



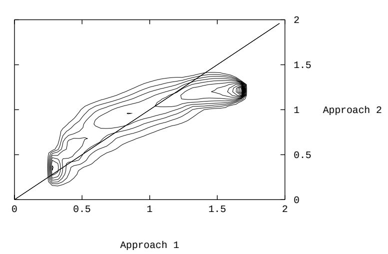
b) 1986–90



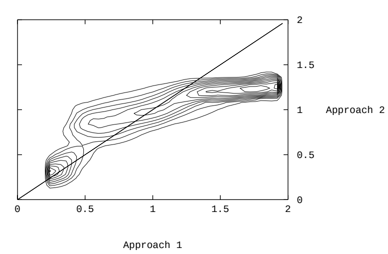
c) 1991



d) 1992–96



e) 1997



f) 1985–97

which figure 4 displays. The positive slope diagonal in every sub-figure represents the continuous counterpart of the diagonals in the transition probability matrices. Consequently, if probability mass abandons such a diagonal, firms will be changing their positions relative to the mean. In particular, if concentration takes place along the negative sloped diagonal, it would indicate that firms are overtaking each other in the efficiency scores' ranking. On the other hand, if probability does not abandon the diagonal, persistence occurs, and firms' positions relative to the mean do not vary. This is the case in 1991 and, more markedly, in the 1992–96 period. However, the trend is not paralleled in 1997 and specially at the beginning of the sample period. It is noticeable, though, that dispersion seems much lower according to the second output definition in 1985 and the 1986–90 period, as probability is far more concentrated along the horizontal axe (approach 2). In 1997, though, the dispersion seems more balanced according to either definition.

6 Conclusions

This study has analyzed the sensitivity of bank cost efficiency scores according to different output measures throughout the 1985–97 period. Efficiency scores have been estimated according to the nonparametric Allocative Data Envelopment Analysis technique, which requires inputs' prices, and two output definitions have been employed, one mostly identified with the asset approach and other considering also savings deposits and other variables, which capture more fully firms' payment and safekeeping services.

Results vary according to either output definition, although time is an issue to account for, as they do not remain constant over the entire sample period. In particular, the asset approach shows a steady increase of simple mean efficiency over the 1985–95 period, although the tendency was completely reversed in the last two sample years. In addition, there has been a convergence process among type of firm's efficiency, as savings banks departed from much lower efficiency scores and now are much the same as commercial banks. However, if savings deposits are considered the tendency is far more stable.

These differing mean values suggest that distributions of efficiency scores could also differ substantially. In order to assess such a question, dispersion indicators might be analyzed. However, estimating nonparametrically—by means of kernel smoothing—the density functions of efficiency scores provided a more accurate view of such distributions. Particularly, the nonparametric approach permitted uncovering all features data might hide, a task in which parametric methods fail. This turns to be of special importance if data are non-normal, or with multiple modes. Results show that, indeed, the shape of the distribution varies much according to either output definition. It is particularly important noticing that several modes exist in many periods, both at high and low parts of the distributions.

But, still, firms' positions relative to the mean according to either output definition must be assessed. Traditionally, this task has been approached by means of correlation coefficients. Our estimations show a sharp increase over the 1985–96 period, partly offset in 1997. But the main contribution constitutes the estimation of transition probability matrices across the output definitions under study and, specially, their continuous counterpart—the stochastic kernels—. They show that, indeed, although results seem somewhat robust in 1991 and 1992–96 periods, firms' relative efficiency scores vary much according to either output measure for the remainder, particularly at the initial sample years. These findings are well represented by the transition probability matrices, but the most accurate view is provided by the joint density functions, conditioned on initial positions. Again, not imposing any functional form allows uncovering all feature data might hide and, in this case, it is possible to precisely identify firms' transition paths.

Consequently, considering different output definitions clearly bias efficiency estimates. These ideas, as stated, were pointed out by several authors. However, perhaps surprisingly, the attention the literature has paid to them is, by far, much lower than that derived from using different techniques, and the gap is particularly important in the case of the study of the Spanish banking firms.[25] In addition, our nonparametric approach to analyze distributions and inter-distribution mobility provides an accurate instrument to evaluate this issues.

Further contributions should attempt to test the robustness of efficiency scores when additional output specifications are considered. This is not straightforward, mostly because of the limitations of our database. Thus, perhaps some efforts should be devoted to create a database capable enough to capture more closely the full range of products and services offered by the banking firm which—and this constitutes an additional issue

[25] As stated, only Grifell-Tatjé et al. (1993) analyzed this issue, finding also noticeable differences. These authors were also surprised by the little attention the literature devoted to this topic. One decade after, contributions are yet to come.

to account for—is varying over time.

References

Aly, H. Y., Grabowski, R., Pasurka, C., and Rangan, N. (1990). Technical, scale, and allocative inefficiencies in U.S. banking: An empirical investigation. *The Review of Economics and Statistics*, 72:211–218.

Andrés, J. and Lamo, A. (1995). Dynamics of the income distribution across OECD countries. Discussion Paper 252, Centre for Economic Performance.

Bauer, P. W., Berger, A. N., Ferrier, G. D., and Humphrey, D. B. (1998). Consistency conditions for regulatory analysis of financial institutions: A comparison of frontier efficiency methods. *Journal of Economics and Business*, 50(2):85–114.

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

Berger, A. N., Hanweck, G. A., and Humphrey, D. B. (1987). Competitive viability in banking: Scale, scope, and product mix economies. *Journal of Monetary Economics*, 20:501–520.

Berger, A. N. and Humphrey, D. B. (1992). Measurement and efficiency issues in commercial banking. In Griliches, Z., editor, *Output Measurement in the Service Sectors*, pages 245–300. The University of Chicago Press, Chicago.

Berger, A. N. and Leusner, J. H. (1997). The efficiency of bank branches. *Journal of Monetary Economics*, 40:141–162.

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

Cao, R., Cuevas, A., and González-Manteiga, W. (1994). A comparative study of several smoothing methods in density estimation. *Computational Statistics and Data Analysis*, 17:153–176.

Clark, J. A. (1996). Economimic cost, scale efficiency, and competitive viability in banking. *Journal of Money, Credit, and Banking*, 28:342–364.

Colwell, R. J. and Davis, E. P. (1992). Output and productivity in banking. *Scandinavian Journal of Economics*, 94:111–129.

Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View*. Wiley, New York.

Favero, C. A. and Papi, L. (1995). Technical efficiency and scale efficiency in the Italian banking sector: A non-parametric approach. *Applied Economics*, 27:385–395.

Ferrier, G. D. and Lovell, C. A. K. (1990). Measuring cost efficiency in banking. *Journal of Econometrics*, 46:229–245.

Grifell-Tatjé, E. and Lovell, C. A. K. (1996). Deregulation and productivity decline:The case of Spanish savings banks. *European Economic Review*, 40(6):1281–1303.

Grifell-Tatjé, E., Prior, D., and Salas, V. (1993). Efficiency scores are sensitive to variable specification: An application to banking. Working paper, Departament d'Economia de l'Empresa, Universitat Autònoma de Barcelona.

Hancock, D. (1985). The financial firm: Production with monetary and nonmonetary goods. *Journal of Political Economy*, 93:859–880.

Hancock, D. (1991). *A Theory of Production for the Financial Firm*. Kluwer Academic Publishers, Boston.

Humphrey, D. B. (1992). Flow versus stock indicators of banking output: Effects on productivity and scale economy measurement. *Journal of Financial Services Research*, 6:115–135.

Hunter, W. C. and Timme, S. (1995). Core deposits and physical capital: A reexamination of bank scale economies and efficiency with quasi-fixed inputs. *Journal of Money, Credit, and Banking*, 27:165–185.

Jones, M. C., Marron, J. S., and Sheather, S. J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, 91(433):401–407.

Klein, M. A. (1971). A theory of the banking firm. *Journal of Money, Credit, and Banking*, 3:261–275.

Kolari, J. and Zardkoohi, A. (1987). *Bank Costs, Structure, and Performance*. Lexington, Massachussets/Toronto.

Kuussaari, H. (1993). Productive efficiency in Finnish local banking during 1985–1990. Discussion Paper 14/93, Bank of Finland.

Kuussaari, H. and Vesala, J. (1995). The efficiency of Finnish banks in producing payment and account transactions. Working Paper 6/95, Bank of Finland–Financial Markets Department.

Lamo, A. (2000). On convergence empirics: Some evidence for Spanish regions. *Investigaciones Económicas*, 24:681–707.

McAllister, P. H. and McManus, D. A. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking and Finance*, 17:389–405.

Mitchell, K. and Onvural, N. M. (1996). Economies of scale and scope at large commercial banks: Evidence from the Fourier flexible functional form. *Journal of Money, Credit, and Banking*, 28:178–199.

Nadaraya, E. A. (1989). *Nonparametric Estimation of Probability Densities and Regression Curves*. Kluwer, Dordrecht.

Park, B. U. and Marron, J. S. (1990). Comparison of data-driven bandwidth selectors. *Journal of the American Statistical Association*, 85(409):66–72.

Park, B. U. and Turlach, B. A. (1992). Practical performance of several data-driven bandwidth selectors. *Computational Statistics*, 7:251–285.

Pérez, F. and Tortosa-Ausina, E. (2000). Product mix clubs, divergence and inequality of spanish banking firms. *Applied Financial Economics*, forthcoming.

Quah, D. T. (1996). Empirics for economic growth and convergence. *European Economic Review*, 40:1353–1375.

Quah, D. T. (1997). Empirics for growth and distribution: Stratification, polarization and convergence clubs. *Journal of Economic Growth*, 2(1):27–59.

Resti, A. (1997). Evaluating the cost-efficiency of the Italian banking system: What can be learned from the joint application of parametric and non-parametric techniques? *Journal of Banking and Finance*, 21:221–250.

Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, New York.

Sealey, C. W. and Lindley, J. T. (1977). Inputs, outputs and a theory of production and cost at depository financial institutions. *Journal of Finance*, 32(4):1251–1266.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Ser.B,53(3):683–690.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

Simonoff, J. S. (1996). *Smoothing Methods in Statistics*. Springer, New York.

Stokey, N. L. and Lucas, R. E. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts.

Wand, M. P. and Jones, M. C. (1994). Multivariate plug-in bandwidth selection. *Computational Statistics*, 9:97–116.

Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing*. Chapman and Hall, London.

Wheelock, D. C. and Wilson, P. W. (2000). New evidence on returns to scale and product mix among US commercial banks. *Journal of Monetary Economics*, forthcoming.