

BANK COST EFFICIENCY AS DISTRIBUTION
DYNAMICS: CONTROLLING FOR
SPECIALIZATION IS IMPORTANT*

Emili Tortosa-Ausina[†]

WP-EC 99-15

Correspondencia: Emili Tortosa-Ausina, Universitat Jaume I
Facultat de Ciències Jurídiques i Econòmiques, Departament d'Economia
Campus del Riu Sec
12080 Castelló (Spain). Tel.: +34 964 728 606, fax.: +34 964 728 591, e-mail: tortosa@uji.es

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
First Edition: November 1999
Depósito Legal: V-4234-1999

IVIE working-papers offer in advance results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

*I must show my gratitude to Fundación Caja Madrid and Fundació Caixa Castelló for financial support, and to the Instituto Valenciano de Investigaciones Económicas, for kindly providing data. The usual disclaimers apply.

[†]Universitat Jaume I

BANK COST EFFICIENCY AS DISTRIBUTION DYNAMICS: CONTROLLING FOR SPECIALIZATION IS IMPORTANT

Emili Tortosa-Ausina

A B S T R A C T

This paper analyzes the dynamics of cost efficiency scores in the Spanish banking industry during the period 1985-1995 and asks how such dynamics are influenced by specialization. Efficiency has been estimated through a nonparametric approach, and a model of distributions dynamics has been applied to assess the evolution of efficiency scores over time. Results show that efficiency at industry level (mean efficiency) has increased importantly and that, considering how the entire distribution evolves, there are some important patterns that mean and standard deviation fail in uncovering, like multi-modality. In addition, controlling for each firm's output mix sheds more light on efficiency dynamics, as efficiency scores tend to become closer much faster.

Key words: Banking, distribution dynamics, X-efficiency, cost efficiency, product mix.

JEL: C14, C30, C61, G21, L5

R E S U M E N

El presente trabajo tiene por objeto el estudio de la dinámica de los índices de eficiencia en la industria bancaria española durante el periodo 1985-1995, así como la manera en que dicha dinámica se ve condicionada por la especialización. La eficiencia se estima a través de un enfoque no paramétrico, y se aplica un modelo de dinámica de las distribuciones para valorar la evolución de los índices de eficiencia en el tiempo. Los resultados indican que la eficiencia a nivel de industria (eficiencia media) se ha incrementado y que, considerando la evolución de la totalidad de la distribución, existen patrones importantes imposibles de detectar por la media y la desviación típica, como la multi-modalidad. Asimismo, condicionar por la especialización de cada empresa arroja más luz sobre la dinámica de la eficiencia, pues los índices tienden a aproximarse entre sí más rápidamente.

Palabras clave: Bancos, dinámica de la distribución, eficiencia-X, eficiencia en costes, especialización.

JEL: C14, C30, C61, G21, L5

1 Introduction

The Spanish banking system has undergone a strong period of change during the last fifteen years, coming mainly from deregulation which, along with other important features like the technological advances or an increasing financial culture are contributing to significantly reshape the industry. This puzzling trends faced by the industry in recent times and the traditional importance of the banking system to the economy as a whole make us consider it as a remarkable topic of policy relevance and research interest.

In particular, one of the most intensively studied topics has been the efficiency of both commercial banks and savings banks. The above stated reasons have driven many researchers to analyze how banks' efficiency is being affected by the new competitive environment. More precisely, it is widely accepted that the removal of entry barriers impels banking firms to decrease their inefficiency levels, in order to increase the industry competitiveness.

However, bank efficiency research studies vary widely on their aims and results, as there are differences regarding the techniques used, the outputs and inputs definitions, the firms being analyzed (savings banks and/or commercial banks), the sample period uncovered, etc. Relative to the technique, some of them consider a nonparametric approach,¹ while others consider an econometric approach.² In some cases the focus lies only on savings banks, essentially due to greater homogeneity of firms and data reliability, while others consider the whole industry, analyzing also commercial banks. In addition, the time scope research studies uncover varies widely, which constitutes a further source of variation.³

But even if a long period is considered, the conclusions drawn on efficiency dynamics use to be based only on two moments of the distribution of the efficiency scores, namely, mean and standard deviation.⁴ These statistics help in acquiring a general picture of the overall efficiency tendencies of the industry, but they do not inform on the differences at firm level. In particular, a time-invariant dispersion may hide very different data structures, which may entail important economic meanings. Some features like multi-modality, non-normality or asymmetry of the data are impossible to uncover by any dispersion indicator.

In this study, a model of explicit distribution dynamics (*medd*) will enable us to fully characterize the dynamics of efficiency scores. And, in particular, due to the tighter competitive conditions firms actually face, an upward tendency in efficiency should emerge. But if, for instance, a pattern like bi-modality persisted (i.e., some firms being more efficient or inefficient than average) and relatively inefficient firms do not abandon the industry, appropriate explanations should be explored.

¹See Doménech (1992), Pastor (1995), Grifell-Tatjé and Lovell (1996, 1997), Prior and Salas (1994), Grifell-Tatjé, Prior and Salas (1992, 1994) or Maudos, Pastor and Pérez (1997).

²See Maudos (1996, 1998), Lozano-Vivas (1997, 1997), Álvarez Cuesta (1998) or Doménech (1993).

³For instance, while Doménech (1992) considers only year 1989 or Grifell-Tatjé, Prior and Salas (1994) consider two years (1989 and 1990), Maudos (1996) uncovers the whole 1985–1994 period.

⁴Álvarez Cuesta (1998) attempts to modelize efficiency dynamics, but he focuses only on savings banks. In addition, our approach to model efficiency dynamics differs widely, as it will be shown below.

Among these, we will consider one which has not been sufficiently stressed in the literature: the role of each firm's output mix. Only the studies by Prior and Salas (1994) and Maudos, Pastor and Pérez (1997) have considered explicitly that different specializations may entail different efficiency levels, as some activities involve more costs than others. Thus, if different product mixes are not controlled for, there could exist an upward bias towards inefficiency. Another way to consider the relevance of output mix consists of defining different banking output specifications, as it involves giving more importance to different specializations. However, despite achieving important differences, this has been done only in Grifell-Tatjé, Prior and Salas (1992) and Tortosa-Ausina (1999a). Our approach to control for specialization when analyzing efficiency and efficiency dynamics will not be the same, as the model considered to assess efficiency scores' dynamics requires a different way to control for some variables.

The study is structured as follows. Section 2 briefly reviews the sources of debate and controversy when measuring efficiency issues, namely, the technique used and the definition of inputs and outputs, estimating the cost efficiency for both savings banks and commercial banks through a DEA approach. Sections 3 and 4 introduce the methodology devoted to the analysis of efficiency dynamics, while section 5 assesses how banks' output mixes may bias efficiency results and efficiency dynamics. Finally section 6 concludes.

2 The study of X-efficiency

Research studies on bank efficiency yield a remarkable dispersion in the results achieved, even when applied to the same database and with similar attempts. The only consensus we may find lies in the dominance of X-inefficiency over scale and product mix ones.⁵ Such a dispersion in results has a twofold source: the technique employed and what we consider banks produce.

2.1 Different techniques to measure efficiency

Parametric and nonparametric methods are the two broad categories where different techniques to measure efficiency lie. In particular, the econometric methods are overwhelmingly used in parametric case, while linear programming techniques are mostly employed within the nonparametric case.

Econometric models estimate efficiency by specifying a functional form, i.e., a cost function (as we are estimating *cost* efficiency) and some assumptions on residuals' distribution.⁶ Observations below the cost frontier would exist only due to random error. On the other hand, linear programming techniques differ widely from the described models. No func-

⁵See Berger and Humphrey (1991).

⁶See Jondrow, Lovell, Materov and Schmidt (1982).

tional form is specified and no observations may lie below the frontier, thus no random error may exist. In addition, these models envelope data more closely, which can be important sometimes.

There exists a clear trade-off between both methodologies: while econometric models specify a functional form which may entail less flexibility or error specification, linear programming techniques fail in decomposing noise and inefficiency. Thus, no methodology dominates the other which makes the choice somewhat arbitrary and more dependent on the aims pursued.⁷

In this study a nonparametric approach has been considered. In particular, the linear programming approach to efficiency analysis, known as DEA (Data Envelopment Analysis), initially developed to compute technical efficiency,⁸ and consequently not requiring the prices of inputs. However, if they are considered, the methodology is not exactly DEA but ADEA (Allocative Data Envelopment Analysis).⁹ Such a methodology considers the specification of a linear programming problem like the following:

$$\begin{aligned} \text{Min}_{x_{js}} \quad & \sum_{j=1}^n \omega_{js} x_{js} \\ \text{s.t.} \quad & y_{is} \leq \sum_{s=1}^S \lambda_s y_{is}, \quad i = 1, \dots, m, \\ & x_{js} \geq \sum_{s=1}^S \lambda_s x_{js}, \quad j = 1, \dots, n, \\ & \lambda_s \geq 0, \quad s = 1, \dots, S, \\ & \sum_{s=1}^S \lambda_s = 1 \end{aligned} \quad (1)$$

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

In order to compute the efficiency scores, the program (1) must be solved for each s firm in the industry. The solution of such a program will be given by the cost minimizing vector x_s^* , given the price vector ω_s and the output vector y_s . Thus, the efficiency score for each s firm is:

$$ES_s = \frac{\omega_s' x_s^*}{\omega_s' x_s} \quad (2)$$

In a similar way, the inefficiency scores are:

$$IS_s = \frac{1}{ES_s} - 1 \quad (3)$$

⁷Some research studies (Resti, 1997) comparing both techniques have come to the conclusion that results do not vary dramatically when applied to the same database, and when this occurs it can be explained by the intrinsic features of each model. Exactly the same occurs in the most famous comparison between both methodologies applied to the banking firm by Ferrier and Lovell (1990), where differences are attributed to the impossibility of decomposing noise from inefficiency, in the nonparametric case, and the imposition of a parametric structure on technology and inefficiency distribution, which fails in decomposing specification error from inefficiency, in the parametric case.

⁸See Banker, Charnes and Cooper (1984).

⁹See Aly, Grabowski, Pasurka and Rangan (1990).

The expression (3) show the amount in which firm s costs would be increased for operating off the efficient frontier.

2.2 A first step in assessing the importance of specialization

In assessing the role of specialization when measuring efficiency issues in the multiproduct firm, a first approach comes precisely from one of the sources of debate and controversy regarding efficiency measurement in banking, namely, the definition of banking output.

We may broadly consider two approaches to this problem: the *intermediation approach* and the *production approach*. The former stresses those issues relating the intermediation labor which banking firms develop, namely, the transmission of funds from agents with financial capacity to investors. On the other hand, the production approach considers banks primarily as services' producers and, consequently, it defines output only from those elements which can be regarded as proxies of such activities.

None approach is satisfactory, as both ignore one or another banking facet. While the intermediation approach does not consider the service producer nature of the bank, the same happens to the production approach relative to its intermediary nature. The desirable choice would be a mix of both approaches, but a sufficient database relative to the production approach is unavailable nowadays, which impels us to consider almost uniquely the intermediation nature of the bank.

However, even within the intermediation approach there exists a further source of controversy, coming from the role of deposits. The reason lies in their twofold nature, as they can be regarded both as inputs and outputs. They are inputs, as their participation is core in the intermediation process. However, they can be also considered as output, due to their payment means features.

Obviously, this debate makes results differ widely when considering different output specifications. However, this variety is not as worrying as the derived from using different techniques in measuring efficiency, as in this case different issues are being measured. If the outputs being considered are not the same according to various output definitions, different specializations are being awarded more or less importance.

According to the above reasonings, a first approach in assessing the importance of product mix could consist of specifying various output definitions. Such an exercise permits assessing how different specializations can bias efficiency results. This has been done in Tortosa-Ausina (1999a), where results achieved for two output specifications vary widely. Such an approach is quite interesting, but in this article the role of specialization and its influence on efficiency measurement and efficiency dynamics is considered in a different way, as it will be shown in section 5.1. With such an approach, the role of specialization will be considered more explicitly.

Thus, only one output definition has been considered. In particular, variables will be given by:

Outputs:

y_1 = fixed income securities+other securities+interbank loans

y_2 = credits to firms and households

Inputs:

x_1 = labor costs

x_2 = savings deposits+other deposits+interbank deposits

x_3 = physical capital

Input prices:

ω_1 = labor costs/number of workers

ω_2 = interest costs/ x_2

ω_3 = cost of capital/ x_3

The output variables are only those within the intermediation approach. The non inclusion of deposits might cause an *a priori* shock to the reader, as most recent studies in banking consider both its input and output nature. However, this has been done on purpose, as such an exercise permits only uncovering the different specializations a bank might exhibit and, as stated, they will be captured differently by means of the approach considered in section 5.1.

2.3 Application to the Spanish banking firms

In order to have a homogeneous database, and regarding the important mergers and acquisitions' process undergone by the Spanish banking industry throughout the sample period, some modifications of the database have been required. Mergers and acquisitions are faced by the literature in different ways. One approach consists of dropping those firms involved in such process; this, however, would entail ignoring some of the largest banks. This has led us to consider a different approach, namely, to backward sum the merged firms, despite this is considered a somewhat controversial approach, as fictitious firms are created. However, it is the only method which allows considering the overwhelming part of the system (more than 90% of gross total assets), ruling out only those firms starting or ending up their operations throughout the sample period.

The application of the linear programming techniques to this homogeneous database yields the results in table 1.¹⁰ A common frontier has been estimated for all firms in the industry (commercial banks and savings banks) according to the output specification considered. Results show a steady increase in efficiency, specially in what savings banks

¹⁰Results both on unweighted means and weighted means are reported. This is done as focusing only on weighted means would imply having the analysis dominated by the largest banks; on the other hand, performing the analysis on an unweighted basis permits extracting meaningful information regardless bank's size.

are concerned and more intense regarding unweighted efficiency. In particular, savings banks start the analyzed period being (on average) much more inefficient than commercial banks (43.93% vs 59.63%) but end it up being more efficient (80.25% vs 77.29%). Such a pattern is the result of a continuous process, with no remarkable ups and downs. Thus, it seems that when considering banking output made up only of earning assets savings banks are becoming as efficient as commercial banks. This result differs substantially from others achieved in previous research studies which, in general, do not find any clear pattern towards efficiency increase or decrease. However, the definition of output followed in our study is very different from those considered in such studies. Such a finding might naturally lead to ask if the source of changes in efficiency might lie in changes in specializations.

3 Conclusions on efficiency dynamics from prior research studies

Although the efficiency of the Spanish banking firms has been an important area of research during the last ten years, even more when considering the strong changes undergone by the industry, no dynamics' insights have been explored explicitly, the exception being Álvarez Cuesta (1998). When conclusions on the evolution of efficiency scores are drawn, this is done by considering only two moments of their distribution: mean and standard deviation (or variance). However, dynamics might be much richer to be precisely captured by such indicators.¹¹ In particular, if both the time and cross-section dimensions of the distribution of the efficiency scores are graphically represented (see figure 1)¹² it is easily noticeable no clear pattern or tendency can be identified.

Which is the overwhelming tendency in figure 1? Are efficiency scores equalizing or, on the other hand, do differences tend to persist? Does any tendency exist at all? Do some firms exhibit special patterns? Obviously, if only mean and standard deviations are being considered, the posed questions cannot be properly answered.

If the rich dynamics figure 1 contains are to be captured, a somewhat more sophisticated technique is required. Although the information provided by mean and standard deviation evolution gives insights on the evolution of industry's efficiency, they fail in identifying some important undergoing changes. In particular, if strong modifications are reshaping the industry, maybe not all firms are reacting equally. If some specializations are changing,¹³ or if some firms are failing in the task of reducing costs, the efficiency scores of such firms could differ widely from industry average. The standard deviation indicator fails in recording these

¹¹It should be pointed out, though, that such studies do not attempt to explicitly model efficiency dynamics.

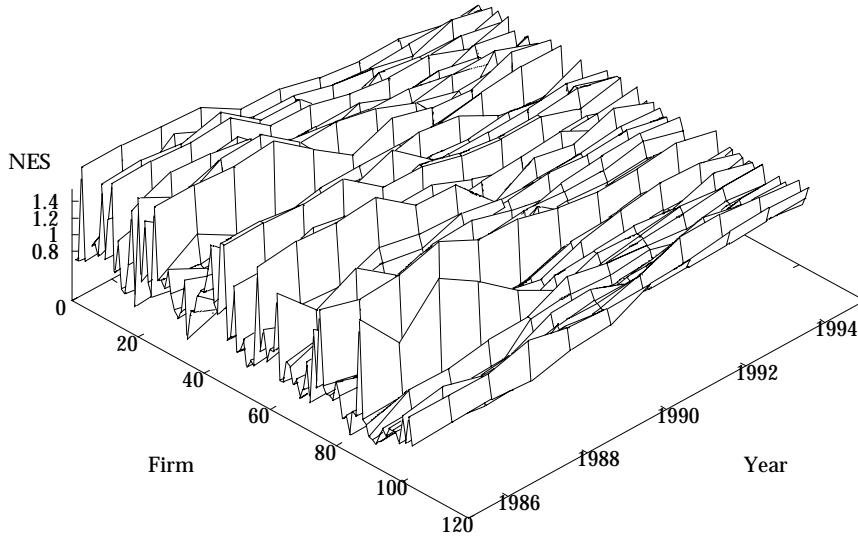
¹²The variable represented is the cost efficiency conditioned by the industry average, as it will be later on more properly explained.

¹³See Pérez and Tortosa-Ausina (1998).

Table 1: Evolution of efficiency, banking firms (1985–1995)

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Commercial banks	Simple mean	59.63	59.81	61.46	59.87	70.60	70.89	73.34	75.51	73.66	69.49
	Weighted mean	82.83	81.87	82.46	84.47	87.80	88.29	85.60	89.95	81.34	81.74
	Standard deviation	21.90	21.70	24.78	24.28	19.70	20.86	19.88	18.53	19.97	22.17
Savings banks	Simple mean	43.93	43.44	43.73	43.26	58.31	59.42	70.54	76.52	77.18	75.63
	Weighted mean	57.19	58.85	60.44	62.55	75.67	77.26	75.34	84.90	82.09	84.67
	Standard deviation	17.80	14.44	17.15	16.79	13.99	13.84	9.85	10.56	10.40	10.92
Total	Simple mean	53.20	52.99	54.19	53.01	65.48	65.99	72.19	75.92	75.11	72.00
	Weighted mean	74.77	74.00	74.93	76.74	83.35	84.25	81.86	88.09	81.61	82.88
	Standard deviation	21.68	20.61	23.58	22.94	18.51	19.00	16.53	15.71	16.73	18.60
											15.13

Figure 1: Efficiency dynamics, 1985–1995



facts as it is unable to capture, for example, if some modes are present in the distribution.

4 An alternative econometric strategy to study efficiency dynamics

To overcome the limitation of drawing conclusions from only two moments of a distribution, it can be used a *model of explicit distribution dynamics* which in turn considers the *entire* distribution.¹⁴ Some of the most important ideas pointing out the advantages of using such a model where initially developed in Quah (1993a, 1993b). Although such studies are primarily focused on the dynamics of per capita income across countries, it is easily applicable to the evolution of efficiency scores in banking. In particular, the limitation Quah attempted to overcome was precisely that derived from extracting conclusions from only two moments of a distribution which, as stated, occurs also in studies of banking efficiency. Obviously, the model goes further in exploring dynamics, as it can properly characterize intra-distribution mobility and long-run behavior.

This model will be based in two elements:

1. The analysis of the cross-section distribution of the variables.

¹⁴See Quah (1996b, 1996c, 1996d, 1997).

2. Modelling of distribution dynamics and its long-run characterization.

Applying this model to the study of banking efficiency or, more properly, to the time evolution of efficiency scores, requires the variables to be slightly modified. In particular, a normalization will be carried out. With this operation the effect on each firm of the variable's behavior for all the industry, which might exhibit generalized oscillations, is corrected. In addition, it is an unavoidable modification if results in this section and section 5 are to be compared.

The new efficiency scores are:

$$NES_s = \frac{ES_s}{\frac{1}{S} \sum_{s=1}^S ES_s} \quad (4)$$

where ES_s are efficiency scores and NES_s are the normalized efficiency scores for all s firms in the sample, $s = 1, \dots, S$. The interpretation is straightforward: if $NES_s = 2$ it would indicate that firm s is twice efficient than the average, while a value of $NES_s = 0.5$ means that it is half efficient.

4.1 Nonparametric density estimation

In order to consider how the *entire* distribution of efficiency scores evolves throughout time, a first step consists of estimating nonparametrically their corresponding density functions at different time periods. The nonparametric approach is desirable and suits perfectly our aims, as it does not impose any *a priori* structure on data, which can be consequently captured. In fact, data might be asymmetric, strongly non-normal or exhibit multiple modes. Although the parametric approach is the most powerful, this phenomena turn out to be invisible to it.¹⁵ Indeed, one of the most important challenges of data analysis consists of uncovering all complexities data might hide and, in an attempt to achieve this, the parametric approach turns out to be definitely unsatisfactory.

The dynamic implications of estimating densities at different points in time are clear: if probability mass tends to be more markedly concentrated around a certain value, convergence would be achieved, namely, (normalized) efficiency scores would tend to equalize. If such a value were 1, the outcome would be a convergence process to the average. Although the opposite outcome (divergence) would imply probability mass being increasingly spread across a wider range, there could exist a wide spectrum of additional results, like different modes emerging or disappearing, phenomenon with strong economic meanings.

However, relying to heavily on the visual aspect of data has undergone strong critiques

¹⁵The nonparametric approach differs widely from the parametric one. The emphasis of the latter consists of, considering a family of parametric density functions $f(\cdot|\theta)$ like the normal $N(\mu, \sigma^2)$, where $\theta = (\mu, \sigma^2)$, obtaining the best estimator $\hat{\theta}$ of θ . However, in the nonparametric case the aim lies in obtaining a good estimator $\hat{f}(\cdot)$ of all the density function $f(\cdot)$. See Scott (1992).

from a historical perspective, although it has some important defenders.¹⁶ The first question that the sceptic may pose is natural: does it make sense going deeply into this analysis if the simple graphic representation allows uncovering any feature inherent to the data? The answer is positive, as when the number of observations increases a graphic representation does not allow to see anything. In order to solve such a shortcoming, data must be *smoothed*, the simplest example of smoothing being the histogram. This is precisely a second argument against the use of nonparametrically estimated density functions: is it not enough using the histogram to uncover data structure? It might be useful as a beginning; in fact, it was the only nonparametric estimator before 1950s. But it faces some important disadvantages,¹⁷ which forces us to select another way to smooth.

To be exact, our analysis will focus on kernel smoothing, which provides a way to uncover data structure without imposing any parametric structure on them. It is not the only method which allows to approach our aims and, following Silverman (1986), it is not the best at any circumstance. However, it is the most widely applicable to many situations, their attributes are easily understood and their discussion permits a better understanding of other methods of density estimation.¹⁸

The kernel method¹⁹ basically consists of estimating the following density function for the variable being analyzed:

$$\hat{f}(x) = \frac{1}{Sh} \sum_{s=1}^S K\left(\frac{x - NES_s}{h}\right) \quad (5)$$

where S is the number of firms, NES_s are the normalized efficiency scores for both output specifications, h is the bandwidth, smoothing parameter or window width and K is a kernel function.²⁰

While kernel's choice determines bumps' shape when plotting function (5), the smoothing parameter h influences it in a different but much stronger manner. If h is too small, an excessive number of *bumps* are generated which do not permit clearly distinguishing data structure. This phenomenon is known as undersmoothing. On the other hand, if h is too large, then oversmoothing occurs and some features inherent to data are hidden. The traditional trade-off between bias and variance which, precisely, depends on the smoothing

¹⁶K. Pearson, for instance.

¹⁷See Silverman (1986).

¹⁸Along with the histogram and the kernel estimator, we may find the naive estimator, the nearest neighbor method, the variable kernel method, the orthogonal series estimators, the maximum penalized likelihood estimators, etc.

¹⁹Some of the most interesting monographs related to this topic are those by Silverman (1986), Scott (1992), Wand and Jones (1995) and Simonoff (1996). More insights can be found in Devroye and Györfi (1985) or Nadaraya (1989).

²⁰Kernel's choice may lie on different alternatives. For instance, Epanechnikov, triangular, Gaussian, rectangular, etc. Given their efficiency use to take values around 90%, such a choice must be based on other arguments, as computing ease. In our case, the Gaussian kernel has been chosen. Anyway, the relevant decision is that regarding the optimal h , which will be considered later on.

parameter, underlies these facts: the larger h , the smaller variance, and vice versa.

Jones, Marron and Sheather (1996) compare different h 's, coming to conclusions stressing the importance of such a question. One of them lies in that, in many occasions, h_{LSCV} (*least squares cross validation*) undersmoothes, while the opposite happens with h_{ROT} (*rules of thumb*).²¹ Second generation methods, though, offer a reasonable balance between these two extremes, which is equivalent to say there is a reasonable balance between bias and variance. This superiority over the first generation methods is being increasingly collected by the literature.²²

These arguments lead to choose the h proposed by Sheather and Jones (1991) from the study by Park and Marron (1990). It is based in the second generation method *solve-the-equation plug-in approach* (h_{SJPI}) and its performance is superior relative to first generation methods, as it has been proved in later studies.²³

4.1.1 Application to the efficiency scores of the Spanish banking firms

The kernel method to nonparametric density estimation has been applied to the efficiency scores computed according to both output measures. In particular, density function (5) has been applied for such scores and the following periods:

1985 and 1995: this allows visually comparing the shape of the cross-section distribution at both the beginning and the end of the regarded period.

1990: as it is the period which divides the considered period into two equally-length sub-periods. Besides, it is a year which witnessed some important transformations in the banking industry, like the fall of the deposits requirements banks must hold in the Bank of Spain, or the climax of the liability's war.²⁴

1986-89 y 1991-94: unavoidable, as we attempt to uncover the entire time dimension of our database.

1985-95: this will enable to make some comparisons between the different clusters of firms which will be considered later on.

Results appear in figure 2, showing the time evolution of the distribution of the efficiency scores for all considered periods. The most important feature is a steady tendency towards efficiency convergence, as probability mass tends to be increasingly more concentrated around unity, despite it seems the process has undergone a slow-down in 1995

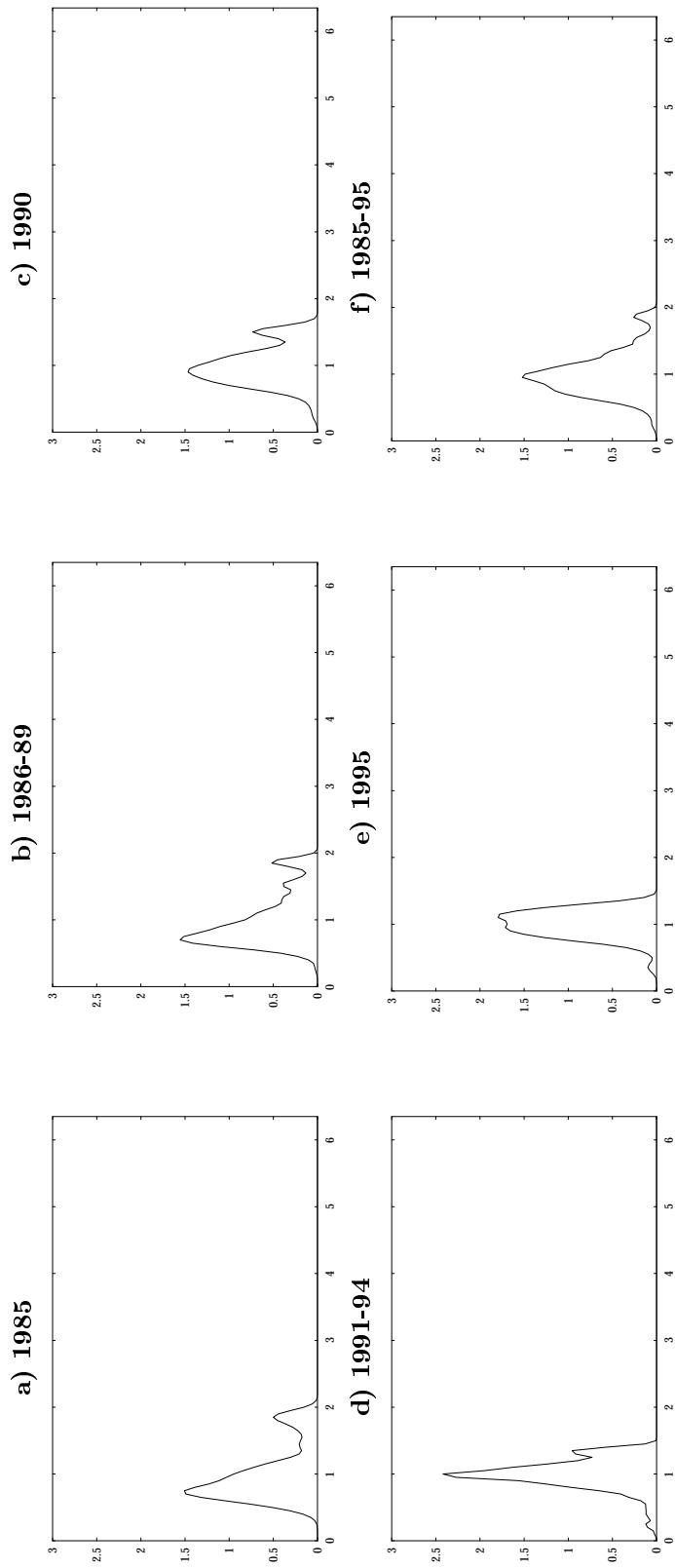
²¹These have been, until now, probably the most applied methods in selecting the optimal bandwidths.

²²See, for instance, simulation studies by Cao, Cuevas and González-Manteiga (1994) or Park and Turlach (1992).

²³Details on their obtaining appear in Sheather and Jones (1991) and Park and Marron (1990). In addition, Steve Marron web's page (<http://www.stat.unc.edu/faculty/marron.html>) provides the Matlab routine which enables their obtaining.

²⁴At the late 1980s there was a sudden increase in the concurrence on savings deposits' interest rates, known as *guerra de las supercuentas*, i.e., *super-accounts' war*.

Figure 2: Normalized efficiency densities (banking firms)



(figure 2.e). The interpretation is straightforward: banking firms possess efficiency scores increasingly closer to industry average, although the features dominating the initial situation (figure 2.a) were not only a much higher dispersion level (standard deviation was 0.407 in 1985 and 0.192 in 1995) but also a very different shape from the final situation (figure 2.e). So, while in 1995 multi-modality is clear, such a phenomenon has almost disappeared in 1995. The initial situation reflected the existence of a group of firms much more efficient than industry average which, as times goes by, are approaching it.

It would be misleading reducing the dynamic analysis of the distribution to only two points in time, because of the important information loss it would entail. If the analysis is spread to the remaining periods, it is noticeable that the situation in 1995 is the result of a very continuous process, with some very efficient firms being increasingly closer to industry average and many others below it which, as time goes by, approach it. In sum, initially less efficient firms are increasingly efficient and vice versa, although the initially more efficient firms approach the average more steadily.

Thus, there has been a transition from relatively high dispersion and multi-modality to another one where the probability mass tends to be more markedly concentrated around unity. Consequently, it is clear that the behavior of the cross-section of the distribution cannot be captured only by mean and standard deviation, requiring a more detailed analysis of the distributions and their mobility throughout time.

4.2 Intra-distribution dynamics and long-run tendencies

Although section 4.1 sheds some light on distribution dynamics, the new approach is still improvable. In particular, despite the dynamic behavior of a distribution might not offer a clear pattern towards convergence or divergence (in the stated way), strong intra-distribution mobility might be undergoing. In other words, strong changes in firms' relative positions might be taking place without being reflected on density functions' shape. In addition, no long-run behavior of the cross-section distribution can be inferred from the results achieved until now. However, inference turns out to be very important when dealing with convergence issues, as this is a concept strongly linked to the idea of limit.

Approaching such questions requires obtaining a *law of motion* of the cross-section distribution within a formal structure, i.e., modelling its dynamics.

4.2.1 The evolving shape of the density functions

Getting insights into such a law and, consequently, drawing conclusions on the variables' dynamic patterns impels us to model²⁵ the stochastic process which takes values that are

²⁵See Quah (1996a, 1996b, 1996c, 1996e, 1997), Andrés and Lamo (1995) or Koopmans and Lamo (1995).

probability measures (λ_t) associated to the cross-section distribution at time t (F_t), where:

$$\forall y \in \mathbb{R} : \lambda_t((-\infty, y]) = F_t(y) \quad (6)$$

Such an aim enables us to build a formal statistical structure which captures the stated phenomena (intra-distribution mobility and long-run behavior). However, the standard econometric analysis does not provide appropriate instruments to model the sequence of distributions' dynamics. With such attempts, we can resort to Markov processes theory and establish a duality in order to approach the problem, so that the same as transition probability functions describe the dynamics of a scalar process, the *stochastic kernels* describe the dynamics or law of motion of a sequence of distributions.²⁶ In other words, the equivalent for distributions of the dynamics of a scalar process is being considered.²⁷

Let λ_t be the probability measure associated to the distribution of each output specification efficiency scores F_t (one for each output specification) at time t , the stochastic kernel²⁸ describing the evolution from λ_t to λ_{t+1} is the mapping M_t to $[0,1]$ of the Cartesian product of efficiency scores and Borel-measurable sets such that.²⁹

$$\forall \text{ set } A \text{ Borel-measurable} : \lambda_{t+1}(A) = \int M_t(y, A) d\lambda_t(y) \quad (7)$$

Notice that the values equation (7) takes are measures or distributions instead of scalars or finite dimensional vectors. Additionally, assuming M_t time-invariant, equation (7) can be re-written as:

$$\lambda_{t+1} = M * \lambda_t \quad (8)$$

where M is a representation of the stochastic kernel encoding information on how starting with a probability measure λ_t associated to the cross-section distribution F_t we end up in λ_{t+1} (associated to F_{t+1}), i.e., on the different firms' relative positions, which is equivalent to knowing partly the dynamics we attempt to model. Thus, estimation of M from the available data allows to empirically quantify distribution dynamics.

Additionally, considering equation (7) and iterating:

$$\lambda_{t+s} = (M * M * \dots * M) * \lambda_t \quad (9)$$

²⁶Stokey and Lucas (1989), secs. 8.1 and 8.3.

²⁷Details on this has been intentionally omitted, due to its excessively technical nature. We have attempted only to provide the ideas such concepts entail. Anyway, we will follow the ideas by Quah (1996b, 1997), Andrés and Lamo (1995), Koopmans and Lamo (1995) and Stokey and Lucas (1989).

²⁸It is hard to completely understand the links between the analysis of distribution dynamics and Markov Processes Theory, in general, and the stochastic kernels, in particular. The study by Durlauf and Quah (1998) is the one which more precisely captures such links which, as stated, turn out to be quite complex.

²⁹See the technical appendix in Tortosa-Ausina (1999b).

This expression allows characterizing (when $s \rightarrow \infty$) the ergodic distribution, thus completely characterizing the efficiency scores' distribution dynamics.³⁰

4.2.2 Application to the Spanish banking firms

Stochastic kernels *inform* on the different analyzed variables' distribution dynamics and their estimation will be based on the nonparametric estimation of bivariate density functions. So, assuming observations refer to each firm position and correspond to a year, changes in firms' relative positions will be analyzed between two years or periods. To be exact, k -year transitions will be analyzed, $k = 1, 11$.³¹

In this (bivariate) case the nonparametric estimation of the density functions will be based again in the kernel method. In the more general multivariate case the function to be estimated adopts the following expression:³²

$$\hat{f}(\mathbf{x}; \mathbf{H}) = S^{-1} \sum_{s=1}^S K_{\mathbf{H}}(\mathbf{x} - \mathbf{NES}_s) \quad (10)$$

where \mathbf{H} is a $d \times d$ matrix (2×2 in our case) known as bandwidth matrix,

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2} \mathbf{x}) \quad (11)$$

and K is a d -variate kernel function.

Estimating a matrix of smoothing parameters \mathbf{H} instead of a unique parameter h entails additional difficulties. If \mathcal{F} is the class of positive definite matrices $d \times d$, then $\mathbf{H} \in \mathcal{F}$. This implies \mathbf{H} has $\frac{1}{2}d(d+1)$ independent entries which, even for small d 's implies considering a high number of smoothing parameters. However, for several reasons we will consider the entries off the main diagonal are zero, so only two bandwidths (one for each coordinate direction) will be considered.³³

The choice of the smoothing parameters in the multivariate case is still in an early stage of development, much earlier than in the univariate case. However, the faced problems when obtaining bandwidths through first generation methods are common both to the univariate and multivariate cases. In this particular bivariate case, the *solve-the-equation plug-in* approach will be considered again. To be exact, the grounds of the analysis will lie in Wand

³⁰The ergodic distribution should not be considered exactly as a prediction of the future, as future realizations of the variables could be influenced by a wide range of ways. This concept should be more properly considered a characterization of last years' tendencies.

³¹In other words, if $k = 1$ then analyzed transitions are between periods 1985 and 1986, 1986 and 1987, 1987 and 1988, etc., whereas if $k = 11$ then only transitions between periods 1985 and 1995 are being considered.

³²A more detailed exposition of the nonparametric estimation of density functions according to kernel smoothing can be found in Wand and Jones (1995) or Simonoff (1996).

³³Despite the strong interest these issues have, space requirements have impelled us to drop the discussion on this topic. However, a longer version of this paper containing more details on it is available from the author upon request.

and Jones (1994) research study. These authors suggest smoothing parameters (different for each coordinate direction) which, in general, perform better than those obtained by means of the *least squares cross validation* method and that have been finally applied when estimating stochastic kernels.

Results on stochastic kernels' estimation by means of bivariate density functions estimated nonparametrically are depicted in figure 3. The top relevance of this analysis would emerge if figure 2 were time-invariant. Such a situation would be fully compatible with changes in firms' relative positions, which only by means of stochastic kernels could be detected.³⁴ This is not our case, as figure 2 exhibits strongly dynamic patterns.

Left-hand side in figure 3 shows firms' mobility (in relative efficiency terms) between periods t and $t+1$ for every two consecutive years in the sample, i.e., changes in firms' relative positions between years 1985 and 1986, 1986 and 1987, 1987 and 1988, and so on, are being considered. Through its analysis and, primarily, through the contour plots, it is noticeable that inter-annual mobility is not very high, as probability mass tends to concentrate along the positively-sloped diagonal in the contour plots, showing persistence in firms' relative efficiency positions. Analogously, if probability concentrated off the positively-sloped diagonal, strong intra-distribution mobility would be undergoing.

However, this picture varies according to the right-hand side in figure 3, plotting transitions for the whole period of analysis (i.e., 11-years transitions).³⁵ In such a case, intra-distribution mobility is high: probability overwhelmingly concentrates off the positively-sloped diagonal, with a narrow range of dispersion in 1995 and a much wider in 1985. Thus, initially more efficient (inefficient) firms than industry's average might end up being as efficient as the initially more inefficient (efficient).

Therefore, although figure 3.a shows persistence in firms' relative positions between two consecutive years by means of probability mass concentrated along the positively-sloped diagonal in the contour plot, results differ regarding 1985–1995 transitions (figure 3.b). In such cases, many firms abandon the positively-sloped diagonal, transiting to other positions and awarding the contour plot a very different shape, with the probability mass strongly concentrated in 1995 and much more spread in 1985.

4.2.3 Long-run tendencies

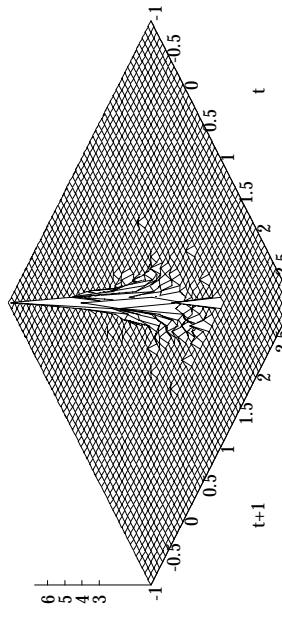
Section's 4.2.2 results allow overcoming one of the limitations of section's 4.1, as changes in firms' relative efficiency positions are identified, by means of intra-distribution movements. However, the long run hypothetic distribution (*ergodic distribution*) is a question still unsolved.

³⁴Those authors familiar to the concepts of σ and β -convergence may find a sort of parallelism here, as the absence of σ -convergence does not preclude the existence of β -convergence.

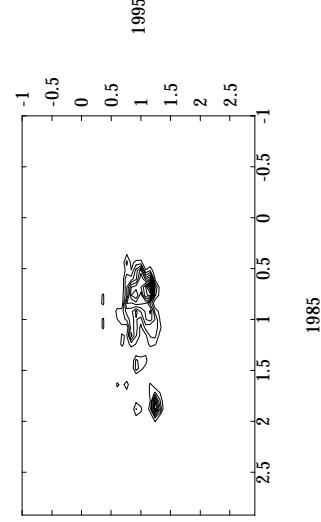
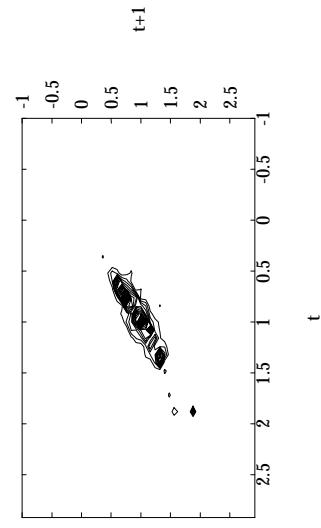
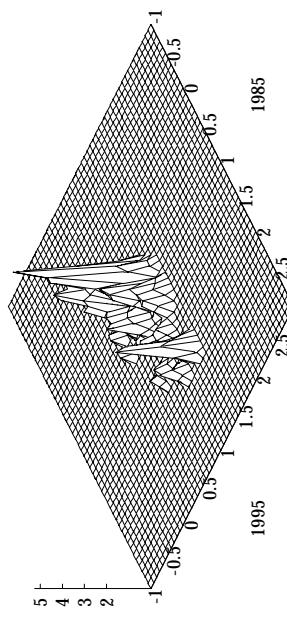
³⁵It would be highly desirable carrying out the same analysis for different transitions, but this involves too much space requirements.

Figure 3: Stochastic kernels, efficiency (banking firms)

a) 1-year transitions



b) 11-years transitions



If the long run is to be characterized by means of estimating the ergodic distribution, the efficiency scores' observations space must be properly discretized. In such a case, measures λ_t turn out to be probability vectors, and the stochastic kernel M becomes a matrix Q . In other words, M and Q are both referring to the stochastic kernels, but in the continuous and discrete case, respectively. Thus, the discrete counterpart to equation (8) is:

$$F_{t+1} = Q \cdot F_t \quad (12)$$

where $Q_{r \times r}$ is a transition probability matrix from one state to another, assuming a finite space of states $E = \{e_1, e_2, \dots, e_r\}$.

The discretization of the observations' space that the analyzed variable (efficiency) might take into r states $e_i, i = 1, \dots, r$ allows clearly interpreting intra-distribution mobility in a way such that state $e_i = (0.5, 3)$ would include those firms which relative efficiency ranks between half and three times industry's average. In addition, cell p_{ij} in matrix $Q_{r \times r}$ would indicate the probability of a firm initially in state i of relative efficiency transiting to state j throughout period or periods (l) being analyzed. In the same way, each row is a probability transitions' vector which helps us in further understanding the analogies to the continuous counterpart: they are equivalent to each probability density for each point in E , when cutting stochastic kernels' figures by a plane parallel to $t + l$ axis.

Results on transition probability matrices are shown in tables 2 and 3. The limits of the states have been chosen in order to have the same amount of probability in each state (20%) in the initial period (1985) which will permit assessing if the ergodic distribution differs substantially from that (uniform) distribution. First column in table 2 shows how many observations (out of 11 years) transited from one state of relative efficiency to another or remained in the same state. So, first value in table 2 showed that, considering all observations for all sample years, 170 firms lied in the first state of relative efficiency. Those firms remained in the same state or transited to other states of relative efficiency in the percentages indicated by each entry in the first row of the matrix, and the same occurs to the remaining four rows in table 2.

If this transition probability matrix were the identity matrix, distributions would be invariant and, in addition, no intra-distribution movements would occur, while if probability tended to be more strongly concentrated off the main diagonal then high intra-distribution mobility would exist.³⁶ The reality, however, is different. For instance, in table 2, the top left-hand entry shows that the most inefficient 20% of banking firms –being less efficient than 67% of the industry's average– remained with relative efficiency scores in that range only with probability 0.62. The remaining 38% transited to other states of higher relative efficiency (18%, 16%, 2% and 2% transited to states 2, 3, 4 and 5 respectively).

Which probability determines a firm ending up in a certain state of relative efficiency?

³⁶According to the limits of the states being considered.

This would be indicated by the ergodic distribution (last row in table 2), showing the highest probability mass concentrated in the fourth state of relative efficiency (35%), i.e., ranging between 0.968 and 1.226 times industry's average efficiency.

Table 2: Unconditioned intra-distribution mobility, banking firms (annual transitions)

NORMALIZED EFFICIENCY					
(Number)	Upper endpoint:				
	0.672	0.767	0.968	1.226	∞
(170)	0.62	0.18	0.16	0.02	0.02
(146)	0.16	0.38	0.39	0.04	0.03
(347)	0.03	0.12	0.55	0.27	0.03
(402)	0.01	0.01	0.22	0.63	0.13
(255)	0.01	0.01	0.05	0.22	0.71
Ergodic distribution	0.07	0.09	0.29	0.35	0.20

Table 3: Unconditioned intra-distribution mobility, banking firms (11-years transitions)

NORMALIZED (EFFICIENCY)					
	Upper endpoint:				
	0.672	0.767	0.968	1.226	∞
	0.00	0.13	0.26	0.61	0.00
	0.00	0.00	0.33	0.63	0.04
	0.08	0.14	0.41	0.32	0.05
	0.05	0.09	0.36	0.36	0.14
	0.06	0.00	0.22	0.33	0.39
Ergodic distribution	0.06	0.09	0.35	0.38	0.12

The picture emerging from table 3 showing transitions between years 1985 and 1995 differs from that in table 2. The 20% most inefficient firms (top-left entry of the matrix) move totally to states 2, 3 and 4 of relative efficiency (accounting for 13%, 26% and 61% of probability mass, respectively). The same occurs to the following 20% of most inefficient firms, abandoning the second state of relative efficiency and moving to the upper-efficiency states (second row in table 3). Thus, the 40% of most inefficient firms, being 77% less efficient than industry's average, have moved totally to states of higher relative efficiency.

On the other hand, those initially most efficient banks show persistence, although at somewhat poor levels (less than 50%). The strong differences disappear when considering long run features, which are similar to those derived from annual transitions: probability mass ends up being overwhelmingly concentrated in states 3 (35%) and 4 (38%).

5 The role of specialization and its influence on efficiency dynamics

Which elements underlie the cost efficiency convergence process? It seems that, according to the output specification considered, many banking firms are becoming increasingly efficient, approaching industry average. Suitable explanations to this have been stated in section 1, and are common to prior research studies of banking efficiency. However, different explanations could be explored. In particular, the role of specialization has not been sufficiently stressed in the literature and, in an environment of major changes, if specializations change,³⁷ efficiency scores could be affected. This point has been forcefully made by Maudos, Pastor and Pérez (1997), who consider that different product mixes entail different efficiency levels, as different products and services require a more intense input use, so that some firms would be mislabelled as inefficient.

Prior research studies on banking efficiency issues do not explicitly consider firms' output mixes as conditioning in measuring efficiency. Such an omission is complete when considering an only output and partial when considering several. And it turns out to be more relevant if what desired is to study the role of specialization as a conditioning in the process of increase or decrease in efficiency inequality. However, does it make sense comparing the efficiency of firms which output mixes differ widely if such output mixes are potential sources of dispersion in efficiency?

A first step in considering how different specializations might bias efficiency scores consists of specifying different output definitions. This has been made in Grifell-Tatjé, Prior and Salas (1992) or Tortosa-Ausina (1999a), coming to relevant conclusions. Thus, it can be interesting and even desirable considering different output definitions, due to the strong debate this topic generates. But we must bear in mind such a debate entails only giving more importance to different specializations. If this is our aim, the techniques used in this article will allow us to approach the problem in a different way which, in addition, turns out to be more comprehensive, as it more properly captures firms' product mixes.

5.1 Product mix clubs in the Spanish banking industry

In order to compare the efficiency of firms with different output mixes, some authors³⁸ have argued separate cost frontiers must be considered for firms focusing on similar activities. Thus, industry is segmented into groups, consequently comparing each firm only to those which specialize in the same products. However, if the number of considered firms is small, the technique used in measuring efficiency (DEA) generates an upward bias in the efficiency scores, simply due to the number of firms being compared.

³⁷See Pérez and Tortosa-Ausina (1998).

³⁸See Maudos, Pastor and Pérez (1997).

To overcome this problem and regarding the different technique being used (*medd*), we will adopt a different approach involving, firstly, considering which firms focus on the same scope of products and services. Thus, a cluster analysis is carried out as, if no *a priori* classification of firms exists according to their output mix similarities, multivariate techniques must be used in order to have unbiased groups or clusters. However, regardless of the variables used and the attempts pursued, studies with these attempts face important problems. In particular, there is no consensus on the number of clusters to consider or how to assess their stability over time.

Relative to the first question, approached in several ways, there is no unique criterion to determine the optimal number of groups. A quick glance to some of the research studies using statistical multivariate techniques to segment the banking industry into groups of firms with similarities would confirm this variety of criteria.³⁹ If we additionally consider the multiplicity of techniques within the multivariate techniques, the variables chosen or the periods being analyzed, the number of groups in the industry may vary widely, as well as their membership and stability over time.

In our case, different considerations⁴⁰ have led us to finally select nine groups⁴¹ which, despite of the problems faced when using these techniques, meet our initial requirements, as they exhibit both within group homogeneity and between group heterogeneity, according to firms' product mixes. In other words, firms in each group are more similar among them regarding their output mixes than compared to firms in other groups.

The other controversial decision in the cluster analysis regards the stability of groups over time. However, despite our groups are stable according to Amel and Rhoades (1988) rule, changes in groups' membership are unavoidable, as the analyzed period has undergone a remarkable process of changes in firms' product mix strategies. Some studies⁴² confirm this assertion, finding several Stable Strategic Time Periods⁴³ throughout the analyzed period.

³⁹See, for instance, Kolari and Zardkoohi (1987, 1995) or Amel and Rhoades (1988) for the case of the American banking industry, or Gual and Hernández (1991) and Sánchez and Sastre (1995) for the Spanish case.

⁴⁰In particular, the variables considered in this study have been eight balance sheet ratios (cash and Bank of Spain, fixed-income securities, interbank loans, credits to firms and households, savings deposits, other deposits, interbank deposits and issued securities) uncovering more than 90% of total assets which more precisely show firms' specializations. More technical issues regard the distance measure employed, in our case the Euclidean squared distance, the method to form clusters, which in our case has been the Ward method, as it minimizes the intra group variance.

⁴¹Full details can be found in Tortosa-Ausina (1999b) and are available from the author upon request. Anyway, this choice meets many of the criteria considered in prior research studies applying cluster analysis to the banking industry.

⁴²See Más (1996).

⁴³See Fiegenbaum and Thomas (1990, 1993) for the definition of a SSTP.

5.2 Does specialization influence tendencies?

Once firms have been satisfactorily clustered into different groups, drawing conclusions on how specialization influences the efficiency convergence process requires a slight modification in the variable of analysis. In particular, while in section 4 efficiency scores were divided by industry's average, in this case they are divided by each firm cluster's average, in the following way:

$$NES_{sk} = \frac{ES_{sk}}{\frac{1}{S_k} \sum_{s=1}^{S_k} ES_{sk}} \quad (13)$$

where NES_{sk} is the efficiency score of firm s affiliated to cluster k relative to its cluster, ES_{sk} is the efficiency score of firm s affiliated to cluster k and S_k is the number of firms in cluster k , $k = 1, \dots, 9$. The interpretation of the new variables of study is different, as we are now controlling for cluster's average instead of industry's average. In this case, if convergence occurs in the sense described above, we will have convergence to cluster's average. In other words, we are not considering *absolute* convergence but *conditional* convergence. This means that there exist other factors apart from industry wide factors being more important in the convergence process, like product mix factors.

5.2.1 Application of the alternative econometric strategy to the new variable of study

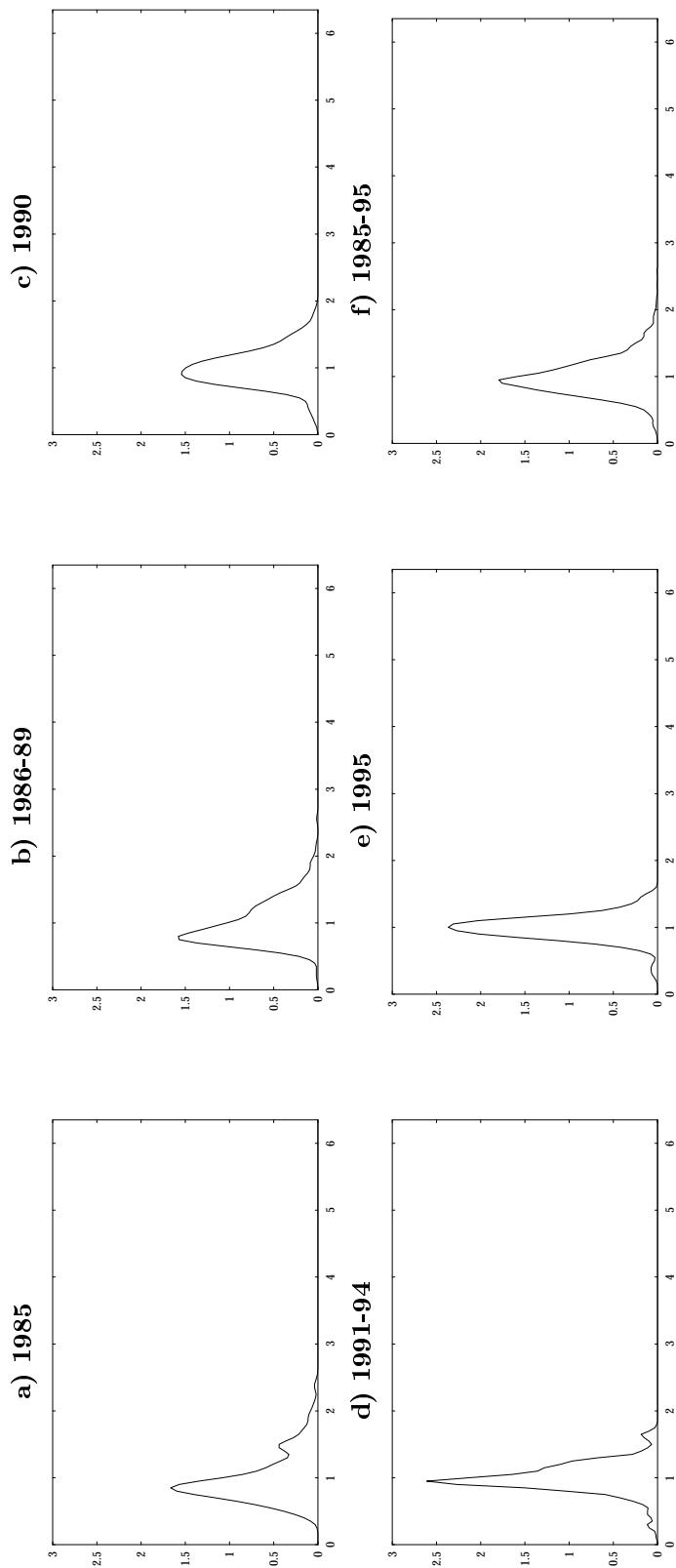
The model of explicit distribution dynamics in section 4 has been applied to the new relative efficiency series (13). If results differed, it would mean the convergence process is influenced not only by industry factors but also by output mix factors.

Figure 4 is the counterpart to figure 2 for product mix conditioned efficiency series. The first conclusion from studying it is that the product mix-relative efficiency scores distributions in figure 2 are tighter and more concentrated than those of industry-relative efficiency scores. Comparing 1985 standard deviations, product mix conditioning gives a reduction of 10.32% over that in industry conditioning efficiency scores. However, the most outstanding feature is that the strong bi-modality of figure 2 almost disappears in its conditioned counterpart, suggesting that efficiency differences tend to diminish when comparing firms with similar output mixes. In addition, although after controlling for specialization some differences still persist, we must bear in mind that the cluster analysis has been done in a 1995 basis, and that specializations have changed remarkably throughout the period.⁴⁴ In this line of results, if the cluster analysis were carried out for each year separately, the conditioned analysis would probably yield much tighter densities for every period.

Figure 5.a is the product mix-conditioned counterpart to figure 3.a, and it shows changes in firms' relative positions between periods t and $t + 1$. Again, high persistence exists, and

⁴⁴See Pérez and Tortosa-Ausina (1998) for further details.

Figure 4: Normalized efficiency densities (product mix conditioned)



somewhat in a higher degree, as probability mass tends to be more markedly concentrated along the positive-sloped diagonal. If figure 5.b is considered, persistence (again) disappears. In this case, which shows changes in firms' relatives positions between years 1985 and 1995, probability tends to abandon the positive sloped diagonal and to be more concentrated along the horizontal axis (1985) than the vertical axis (1995). Consequently, firms' relative positions are less dispersed in 1995, and this occurs in a higher degree than in the industry conditioned counterpart figure (see figure 3.b). However, it is not clear that strong dynamic patterns should exist, as in this product mix-conditioned case firms' initial relative efficiency scores are less dispersed.

Tables 4 and 5 are the discrete counterparts (transition probability matrices) to the stochastic kernels in figure 3. Results show that probability tends to concentrate more markedly in a unique relative efficiency state. In particular, annual transitions show 40% of probability mass concentrates in state 4, while if 11-years transitions are considered, this pattern is more accentuated (49%). On the other hand, probability in state 1 of relative efficiency is only 8% in the annual transitions case and 6% in the 11-years transitions case. This, again, turns out to be of a paramount importance, as probability in 1985 was uniformly distributed (20%) across the five considered states.

Table 4: Conditional (product mix) intra-distribution mobility, banking firms (annual transitions)

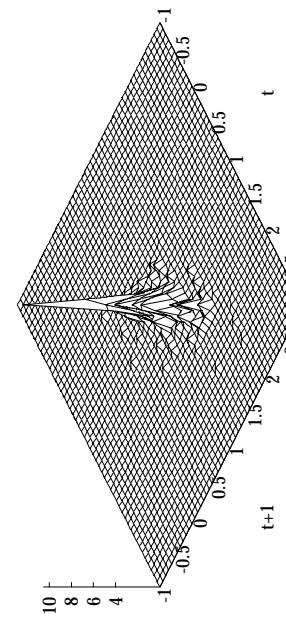
NORMALIZED EFFICIENCY					
(Number)	Upper endpoint:				
	0.727	0.846	0.964	1.236	∞
(154)	0.52	0.25	0.13	0.09	0.01
(206)	0.15	0.42	0.30	0.13	0.00
(293)	0.05	0.19	0.46	0.29	0.01
(467)	0.01	0.04	0.18	0.67	0.10
(199)	0.02	0.03	0.00	0.26	0.69
Ergodic distribution	0.08	0.14	0.23	0.40	0.15

Table 5: Conditional (product mix) intra-distribution mobility, banking firms (annual transitions) (11-years transitions)

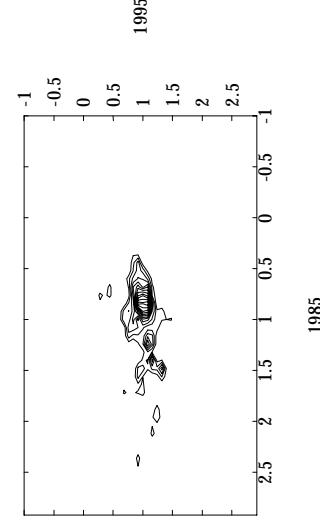
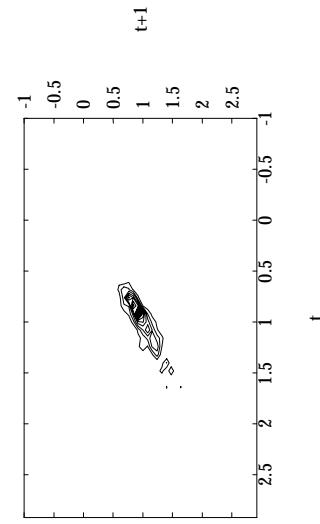
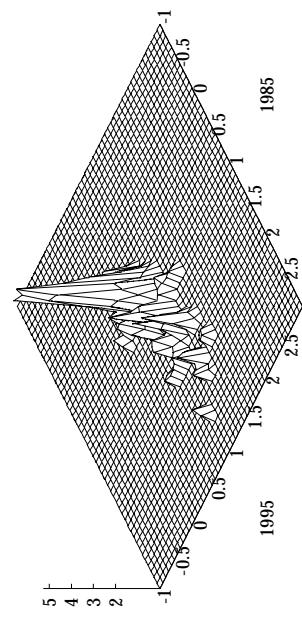
NORMALIZED EFFICIENCY					
	Upper endpoint:				
	0.727	0.846	0.964	1.236	∞
	0.05	0.09	0.48	0.38	0.00
	0.04	0.13	0.33	0.50	0.00
	0.00	0.05	0.25	0.65	0.05
	0.09	0.26	0.13	0.43	0.09
	0.05	0.00	0.19	0.52	0.24
Ergodic distribution	0.06	0.17	0.21	0.49	0.07

Figure 5: Stochastic kernels, efficiency (product mix conditioned)

a) 1-year transitions



b) 11-years transitions



6 Final remarks

The analysis of X-efficiency in the Spanish banking industry has been studied thoroughly during the last few years. Different reasons motivated the research studies, but primarily those regarding the reshaping of the industry due to a strong deregulation process, important technological advances, significant financial market innovation, a growing financial culture or an increasing internationalization of banking firms.

Results vary across studies, as techniques to estimate efficiency, concepts of what banks produce and firms considered (commercial banks and/or savings banks) or even the issues being measured differ widely. However, studies have not explicitly considered how efficiency scores vary over time at firm level, as conclusions use to be industry wide. In addition, it has not been fully assessed how different concepts of what banks produce might bias results. Only Grifell-Tatjé, Prior and Salas (1992) did this, finding important differences.

In this study a nonparametric approach has been considered to estimate Spanish banking firms' cost efficiency throughout period 1985–1995. Results show mean efficiency, according to the output specification being considered, has grown considerably, particularly in what savings banks are concerned. This conclusions have been drawn on an industry basis. But differences at firm level may be very important, as if significant differences persist, industry structure might be affected. However, if individual differences are to be assessed, a somewhat different view is required, and bearing this in mind a *model of explicit distribution dynamics* has been considered. This model focuses on how distributions (in this case, the distribution of efficiency scores) evolve over time and, according to this, which might be the (probable) long run distribution. Its application permits uncovering some features of the distribution hidden to mean and standard deviation, like a strong bi-modality at the beginning of the period which has been reducing over time, and that the decrease of efficiency inequalities has undergone a slow down in 1995.

In addition, it has been also considered how different specializations might bias efficiency dynamics. This has been done by means of multivariate statistical techniques which cluster together those firms with similar output mixes.⁴⁵ According to this, when conditioning each firm's efficiency score on its cluster's average, the convergence process is much stronger, as probability tends to concentrate much faster around the unity. This constitutes an important explanation to why differences in cost efficiency at firm level may persist, with the relatively inefficient firms not abandoning the industry: different output mixes entail some firms being more costly but, once specializations have been controlled for, not more inefficient.

The picture emerging is an industry where cost efficiency levels are increasing both at industry and firm level, but this pattern is significantly influenced by each firm's choice of what products and services to focus on. Once this has been controlled for, most of the

⁴⁵A different approach is followed by Tortosa-Ausina (1999a), by specifying two output definitions.

remaining differences in efficiency scores are removed.

References

- Álvarez Cuesta, R. (1998). Eficiencia técnica de las cajas de ahorros. *Revista de Economía Aplicada*, 6(16):179–191.
- Aly, H.Y., Grabowski, R., Pasurka, C. and Rangan, N. (1990). Main Patterns of Economic Growth in OECD Countries. *The Review of Economics and Statistics*, 72:211–218.
- Amel, D.F. and Rhoades, S.A. (1988). Strategic groups in banking. *The Review of Economics and Statistics*, 70:685–689.
- Andrés, J. and Lamo, A. (1995). Dynamics of the income distribution across OECD countries. Discussion Paper 252, Centre for Economic Performance.
- Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, 30:1078–1092.
- Berger, A.N. and Humphrey, D.B. (1991). The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics*, 28:117–148.
- 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.
- Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation: The L₁ View*. Wiley, New York.
- Doménech, R. (1992). Medidas no paramétricas de eficiencia en el sector bancario español. *Revista Española de Economía*, 9:171–196.
- Doménech, R. (1993). Funciones de costes para la banca española: un análisis con datos de panel. *Investigaciones Económicas*, 17:263–284.
- Durlauf, S.N. and Quah, D.T. (1998). The new empirics of Economic Growth. Discussion Paper 384, Centre for Economic Performance.
- Ferrier, G.D. and Lovell, C.A.K. (1990). Measuring cost efficiency in banking. *Journal of Econometrics*, 46:229–245.
- Fiegenbaum, A. and Thomas, H. (1990). Strategic groups and performance; the U.S. insurance industry, 1970–84. *Strategic Management Journal*, 11(3):197–217.
- Fiegenbaum, A. and Thomas, H. (1993). Industry and strategic group dynamics: Competitive strategy in the insurance industry, 1970–1984. *Journal of Management Studies*, 30(1):69–105.

- 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. and Lovell, C.A.K. (1997). The sources of productivity change in Spanish banking. *European Journal of Operational Research*, 98:365–381.
- Grifell-Tatjé, E., Prior, D. and Salas, V. (1992). Eficiencia frontera y productividad en las cajas de ahorros españolas. Documento de Trabajo 92–1992, Fundación FIES.
- Grifell-Tatjé, E., Prior, D. and Salas, V. (1994). Eficiencia de empresa y eficiencia de planta en los modelos frontera no paramétricos. Una aplicación a las cajas de ahorro españolas. *Revista Española de Economía*, 11(1):139–159.
- Gual, J. and Hernández, A. (1991). Costes operativos, tamaño y especialización en las cajas de ahorro españolas. *Investigaciones Económicas*, 15(3):701–726.
- Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 34:335–348.
- 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.
- Kolari, J. and Zardkoohi, A. (1987). *Bank Costs, Structure and Performance*. Lexington, Massachussets/Toronto.
- Kolari, J. and Zardkoohi, A. (1995). Economies of scale and scope in commercial banks with different output mixes. Working paper, Texas A&M.
- Koopmans, R. and Lamo, A. (1995). Cross-sectional firm dynamics: Theory and empirical results from the chemical sector. Discussion Paper 229, Centre for Economic Performance.
- Lozano-Vivas, A. (1997). Profit efficiency for Spanish savings banks. *European Journal of Operational Research*, 98:382–395.
- Más, F.J. (1996). *Competencia y dinámica de grupos estratégicos. Aplicación al sector bancario español*. Universidad de Alicante, Alicante.
- Maudos, J. (1996). Eficiencia, cambio técnico y productividad en el sector bancario español: una aproximación de frontera estocástica. *Investigaciones Económicas*, 20(3):339–358.
- Maudos, J. (1998). Market structure and performance in Spanish banking using a direct measure of efficiency. *Applied Financial Economics*, 8:191–200.
- Maudos, J., Pastor, J.M. and Pérez, F. (1997). Competencia y evolución de la eficiencia en el sector bancario español: la importancia de la especialización. Unpublished manuscript paper.

- 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.
- Pastor, J.M. (1995). Eficiencia, cambio productivo y cambio técnico en los bancos and cajas de ahorro españolas: un análisis de la frontera no paramétrico. *Revista Española de Economía*, 12(1):35–73.
- Pérez, F. and Tortosa-Ausina, E. (1998). Product mix of the Spanish banking firms: Do competition clubs exist? Working Paper-Serie EC 98–02, IVIE.
- Prior, D. and Salas, V. (1994). Eficiencia técnica de las cajas de ahorros españolas y sus factores determinantes. *Papeles de Economía Española*, 58:141–161.
- Quah, D.T. (1993a). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37:426–434.
- Quah, D.T. (1993b). Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics*, 95(4):427–443.
- Quah, D.T. (1996a). Convergence as distribution dynamics (with or without growth). Discussion Paper 317, CEPR.
- Quah, D.T. (1996b). Convergence empirics across economies with (some) capital mobility. *Journal of Economic Growth*, 1(1):95–124.
- Quah, D.T. (1996c). Empirics for economic growth and convergence. *European Economic Review*, 40:1353–1375.
- Quah, D.T. (1996d). Regional convergence clusters across Europe. *European Economic Review*, 40(3–5):951–958.
- Quah, D.T. (1996e). Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal*, 106(437):1045–1055.
- 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.

- Sánchez, J.M. and Sastre, M.T. (1995). ¿Es el tamaño un factor explicativo de las diferencias entre entidades bancarias? Documento de trabajo 9512, Banco de España, Servicio de estudios.
- Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, New York.
- 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 Jr, R.E. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts.
- Tortosa-Ausina, E. (1999a). Convergence in efficiency of the Spanish banking firms as distribution dynamics. Working Paper-Serie EC 99-14, IVIE.
- Tortosa-Ausina, E. (1999b). *Especialización productiva, eficiencia y convergencia de las empresas bancarias españolas*. PhD thesis, Universitat Jaume I.
- 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.