Centre for International Capital Markets

Discussion Papers ISSN 1749-3412

Identification of Segments of European Banks with a Latent Class Frontier Model

Carlos Pestana Barros, Guglielmo Maria Caporale and Luis A. Gil-Alana

No 2007-12

IDENTIFICATION OF SEGMENTS OF EUROPEAN BANKS WITH A LATENT CLASS FRONTIER MODEL

Carlos Pestana Barros[a]
Guglielmo Maria Caporale[b*]
Luis A. Gil-Alana[c]

[a] Instituto Superior de Economia e Gestão (ISEG), Technical University of Lisbon
[b]Centre for Empirical Finance, Brunel University, London
[c] University of Navarra

September 2007

* Corresponding author. Email: Guglielmo-Maria.Caporale@brunel.ac.uk

Abstract

This paper analyses technical efficiency of European banks over the period 1996-2003 with unbalanced panel data techniques. A latent class frontier model is used which allows the identification of different segments in the production frontier. We find that there are three statistically significant segments in the sample. Therefore, we conclude that no common banking policy can be effective for all the banks included in the sample, and that banking policies by segments are required instead.

Keywords: European banking; latent class frontier model; technical efficiency.

JEL Classification: C23, G21

1. Introduction

This paper applies a parametric frontier model to the European banking industry. Over the last decade or so a vast literature has emerged, which uses a variety of approaches to analyse productivity in banking. Examples are the studies of Casu et al. (2004), Williams (2001), Kumbhakar et al. (2001), Bauer et al. (1993), Humphrey and Pulley (1997), Stiroh (2000), Alam (2001) and Berger and Mester (2003). Only a few papers, though, estimate a latent frontier model – one of the few exceptions is the study due to Orea and Kumbakhar (2005), who analyse the efficiency of Spanish banks. The key advantage of this approach is that it enables one to define endogenous bank segments in the sample under consideration.

The present paper extends their research to several European countries. The motivation is the following. First, for most of their lifetime, European saving banks have operated within a well-defined market structure and in accordance with a clear regulatory policy. However, in recent years, at European level the market structure of the banking sector has evolved as a result of the European Union's Single Market Programme (SMP), which was established in 1992 with the aim of facilitating the free movement of goods and services across member states in order to improve efficiency. The consolidation process that has taken place throughout Europe has also resulted in the number of banking firms falling from 9,100 in 1997 to 7,500 by 2003 [ECB 2004], and one would expect these structural changes also to have an impact on the behaviour and subsequent productivity performance of banks. The expected result of these changes is an increase in competition, and hence an improvement in efficiency. Second, a regulatory policy in each European state has been in existence in all national European markets since 1999. Competition and regulation are related to efficiency in the European banking industry – see Casu et al. (2004), Williams (2001), Kumbhakar et al.

(2001) and Orea and Kumbhakar (2005). Lastly, although European banks are clustered by countries, it is also possible for banks from different countries to be in the same cluster, which clearly has policy implications.

Our principal aim is to endogenously identify clusters of banks in Europe. This is possible using our chosen framework, which differs from the traditional approaches assuming homogeneity of all banks. Moreover, this method is more effective than standard regression models in simultaneously segmenting and profiling banks in a sample. If the characteristics identifying banks are known, then cluster policy implications can be derived.

The paper is organised as follows: Section 2 provides a brief overview of the existing literature on European banking efficiency and of the main features of the banking sector in Europe. Section 3 outlines the econometric model. The underlying theoretical model and the hypotheses to be tested are discussed in Section 4. Section 5 presents the empirical results. Some concluding remarks follow in Section 6.

2. The European Banking Sector

Papers on banking using non-parametric models rely either on DEA (Data Envelopment Analysis) or stochastic frontier models. Two recent extensive surveys on frontier models applied to banking are Molyneux et al. (2001) and Berger and Humphrey (1997), who summarised all the research done in the area until 1997.

Recent papers using DEA not cited in the above surveys include Dietsch and Weill (2000), who focused on 661 commercial, mutual and saving banks from 11 EU countries for the period 1992 to 1996. They estimated a DEA model, the Malmquist index, and a profit frontier model, and found an increase in both cost and profit efficiency, as well as in total productivity, mainly due to positive technical progress. Garden and Ralston (1999) estimated the technical and allocative efficiencies of

Australian credit unions. Worthington (1999) applied the DEA in a two-stage procedure to analyse Australian credit unions. Chen and Yeh (2000) assessed the technical efficiency of Taiwanese banks. Drake (2001) used the DEA to estimate the Malmquist index for UK banking. Casu et al. (2004), Williams (2001) and Molyneux and Williams (2005) studied various European countries.

A recent paper using stochastic frontier models is Bos and Schmiedel (2007) who estimate cost and profit met-frontiers for the European banking sector. Valverde, Humphrey and Paso (2007) analyse the efficiency of Spanish banks with parametric and stochastic frontier models, and Dietsch and Weill (2000) compare the efficiency of French and Spanish banking with a parametric-free distribution approach. Latent frontier models have been applied in banking so far only by Orea and Kumbhakar (2004).

We use a dataset on European banks from 1996 to 2003. There are 5,721 saving banks, 6,180 commercial banks and 11,816 cooperative banks in the sample. Table 1 shows the distribution of savings banks across countries over the period examined.

Table 1: Number of Saving Banks & Observations: by Country, 1996-2003

Country	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*Total*
Austria	13	65	66	67	67	67	65	65	475
Belgium	15	14	14	13	12	11	10	7	96
Finland	1	1	1	1	1	1	1	1	8
France	22	22	22	29	30	30	30	28	213
Germany	592	584	588	572	556	527	490	233	4.142
Ireland	3	3	3	3	3	2	2	2	21
Italy	58	58	58	57	54	56	53	13	407
Luxembourg	1	1	1	1	1	1	1	1	8
Portugal	3	3	3	3	3	3	3	2	23
Spain	30	38	39	39	39	41	42	46	314
All banks	741	792	798	787	767	740	698	398	5.721

Savings banks are found to be most common in Germany, Austria and Italy.

Table 2 reports the number of the commercial banks in Europe by country in the period analysed.

Table 2: Number of Commercial banks: by Country, 1996-2003

Countries	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*Total*
Austria	34	35	40	41	46	48	48	36	328
Belgium	40	42	35	33	32	32	28	17	259
Finland	5	7	7	7	7	6	5	5	49
France	186	188	181	172	165	166	140	107	1.305
Germany	182	192	187	177	180	180	168	130	1.396
Greece	10	16	14	14	14	14	16	13	111
Ireland	16	19	22	24	25	28	30	25	189
Italy	95	110	108	117	110	118	112	46	816
Luxembourg	114	115	108	114	106	93	88	63	801
Netherlands	36	34	31	31	29	32	34	24	251
Portugal	21	22	22	22	18	15	15	8	143
Spain	72	75	72	65	63	66	63	56	532
All banks	811	855	827	817	795	798	747	530	6.180

It can be seen that Germany and France have the highest number of commercial banks, whilst the lowest is found in Finland.

Table 3 presents the corresponding figures for cooperative banks.

Table 3: Number of Cooperative Banks & Observations: by Country, 1996-2003

Countries	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*Total*
Austria	19	18	26	31	45	50	40	20	249
Belgium	10	9	8	7	7	7	7	6	61
Finland	1	1	1	1	1	1	1	1	8
France	64	67	70	101	106	104	91	83	686
Germany	1.012	993	1.184	1.171	1.069	974	816	370	7.589
Italy	181	428	417	475	482	500	477	132	3.092
Luxembourg	2	2	2	2	2	2	2	2	16
Netherlands	2	2	1	1	1	1	1	1	10
Portugal	1	1	1	1	2	2	2	1	11
Spain	10	15	13	8	11	15	13	7	92
All banks	1.303	1.537	1.723	1.798	1.726	1.656	1.450	623	11.816

Clearly, Germany and Italy dominate, with Finland and Portugal displaying the smallest number.

3. Latent Class Frontier Models

The theory underlying the model of bank clusters adopted in this paper was developed by Porter (2000). His model explains the persistence of geographical agglomeration in terms of population agglomeration and knowledge embodied in human capital and acquired through experience. A bank cluster is a geographic concentration of interconnected banks in a market. Agglomeration reflects population density and knowledge increasing growth. The clustering indicates that banks are organizations attempting to maximize profits, subject to resource constraints (Varian,1987).

In our empirical analysis, we follow a stochastic frontier approach. This came into prominence in the late 1970s as a result of the work of Aigner, Lovell and Schmidt (1977), Battese and Corra (1977) and Meeusen and Van den Broeck (1977). In this framework it is assumed that the residuals have two components (noise and inefficiency). The frontier is estimated using maximum likelihood techniques, and the residuals represent the difference between the observations and the frontier. A stochastic cost function can be written as:

$$\ln C_{it} = C(X_{it}) + \varepsilon_{it}; \quad \varepsilon_{\mathrm{it}} = v_{it} + u_{it}\,; \quad \mathrm{i} = 1,2,\ldots\mathrm{N}, \quad \mathrm{t} = 1,2,\ldots T \tag{1}$$

where C_{it} represents a scalar cost of the decision-unit i under analysis in the t-th period; X_{it} is a vector of variables including input prices and output; and ε is the error term. The symmetric component, v, captures statistical noise and it is assumed to follow a distribution centered at zero, while u is a non-negative term that reflects technical inefficiency and it is usually assumed to follow a one-sided distribution. The two components v and u are assumed to be independent of each other.

Given that estimation of equation (1) yields merely the residual ε, rather than the inefficiency term u, this term in the model must be calculated indirectly. In the case of panel data, such as those used in this paper, Battese and Coelli (1988) use the conditional expectation of u_{it}, conditioned on the realized value of ε, as an estimator of u_{it}. In other words, $E\left[u_{it} / \varepsilon_{it}\right]$ is the inefficiency for the i-th bank at time t. Following Greene (2001), we can write equation (1) as a latent class model:

$$\ln C_{it}\big|_j = f(x_{it})\big|_j + v_{it}\big|_j + u_{it}\big|_j \,, \tag{2}$$

where subscript i denotes the firm, t indicates time and j represents the different classes. It is assumed that each club belongs to the same group in all periods.

Assuming that v is normally distributed and u follows a half-normal distribution, the likelihood function (LF) for each club i at time t for group j is (see Greene, 2005):

$$LF_{ijt} = f\left(C_{it} \middle| x_{it}, \beta_j, \sigma_j, \lambda_j\right) = \frac{\Phi\left(\lambda_j \cdot \varepsilon_{it|j} / \sigma_j\right)}{\Phi(0)} \cdot \frac{1}{\sigma_j} \cdot \phi\left(\frac{\varepsilon_{it|j}}{\sigma_j}\right), \tag{3}$$

where $\varepsilon_{it|j} = ln\, C_{it|j} - \beta'_j x_{it}$, $\sigma_j = \left[\sigma_{uj}^2 + \sigma_{vj}^2\right]^{1/2}$, $\lambda_j = \sigma_{uj} / \sigma_{vj}$, and ϕ and Φ denote the standard normal density and cumulative distribution function respectively. The likelihood function for club i in group j is obtained as the product of the likelihood functions in each period:

$$LF_{ij} = \prod_{t=1}^{T} LF_{ijt} \,. \qquad . \tag{4}$$

The likelihood function for each bank is obtained as a weighted average of its likelihood function for each group j, using as weights the prior probabilities of class j membership:

$$LF_i = \sum_{j=1}^{J} P_{ij} LF_{ij} \ . \tag{5}$$

The prior probabilities must be in the unit interval: $0 \le P_{ij} \le 1$. Furthermore, the sum of these probabilities for each group must be one: $\sum_j P_{ij} = 1$. In order to satisfy these two conditions we parameterised these probabilities as a multinational logit. That is:

$$P_{ij} = \frac{\exp(\delta_j q_i)}{\sum_{j=1}^{J} \exp(\delta_j q_i)} , \tag{6}$$

where q_i is a vector of variables which are used to split the sample, and δ_j is the vector of parameters to be estimated. One group is chosen as the reference in the multinational logit. The overall log-likelihood function is obtained as the sum of the individual log-likelihood functions:

$$\log LF = \sum_{i=1}^{N} \log LF_i = \sum_{i=1}^{N} \log \sum_{j=1}^{J} P_{ij} \prod_{t=1}^{T} LF_{ijt} \ . \tag{7}$$

The log-likelihood function can be maximised with respect to the parameter set $\theta_j = (\beta_j, \sigma_j, \lambda_j, \delta_j)^T$ using conventional methods (Greene, 2005). Furthermore, the estimated parameters can be used to estimate the posterior probabilities of class membership using Bayes Theorem:

$$P(j/i) = \frac{P_{ij} LF_{ij}}{\sum_{j=1}^{J} P_{ij} LF_{ij}} \ . \tag{8}$$

4. Theoretical Framework and Hypotheses of Interest

We estimate a latent frontier model to analyze the efficiency of banks in several European countries. The underlying economic theory is given by Porter’s (2000) cluster model. This model was proposed for industrial economics, but can be adapted to explain banking clustering in a geographical area. According to this theory the persistence of geographical agglomeration is explained by population agglomeration and knowledge embodied in human capital and acquired through experience. A bank cluster is a geographic concentration of interconnected banks in a region. This agglomeration derives from population density and knowledge. The clustering of banks in a global economy lies increasingly in local characteristics.

Consider the banks operating in the European market. The frontier model allows us to test the following null hypotheses (see Barros, Ferreira and Williams, 2007):

Hypothesis 1 (Savings banks): Savings banks perform efficiently searching for profits. This hypothesis is based on previous research on banking (Williams, Peypoch and Barros, 2007) and on strategic-group theory (Caves and Porter, 1977) which explains differences in efficiency scores in terms of differences in the structural characteristics of units within an industry. In the case of saving banks, units with similar asset configurations pursue similar strategies with similar performance results (Porter, 1979). Although there are different strategic options in the different sectors of an industry, because of mobility impediments, not all options are available to each bank, causing a spread in the efficiency scores of the industry. Therefore it is assumed that savings banks adopt cluster-specific efficient strategies.

Hypothesis 2 (Commercial banks): Commercial banks perform efficiently searching for profits (Peypoch, Barros and Williams, 2007). This hypothesis is based on the theory of transaction costs and property rights, as in Klein, Crawford and Alchian (1978), Williamson (1979, 1985) and Gross and Hart (1986). There are two critical assumptions: First, the firms cannot write complete contracts concerning their funds allocation based on the interest rate. Second, investments are specific to banks' assets so that the same investment is less valuable with different assets. When both assumptions hold, the theory predicts that firms under-invest because they are afraid that their relationship with the other firms may end at same point. To minimise under-investment, firms allocate dedicated asset specificity (Williamson and Joskow 1985), which refers to investment which takes place with the prospect of selling a significant amount of product to a particular customer. Given this asset-specific strategy, commercial banks are therefore assumed to be efficient.

Hypothesis 3 (cooperative banks): Cooperative banks perform efficiently searching for profits (Barros, Peypoch and Williams, 2007). This hypothesis is based on previous research on cooperative banking at European level (Molyneux and Williams, 2005).

Each of these hypotheses will be tested with the latent frontier model.

5. Empirical Analysis

Financial statement data for commercial banks operating in fifteen EU countries between 1996 to 2003 have been obtained from the BankScope database. Table 4 reports some descriptive statistics.

Table 4: Variable definitions

Variable	Description	Minimum	Maximum	Mean	Std. Dev
Log Cost	Logarithm of total operational cost	11.395	7.638	15.888	1.493
Log PL	Logarithm of the price of labour measured dividing the total wages by the number of equivalent workers	2.817	0.782	6.409	0.396
Log PK1	Logarithm of the price of capital-Stock, measured dividing the stock by the assets	2.342	4.240	3.317	2.715
LogPK2	Logarithm of the price of capital-premises, measured dividing the premises by the total assets	1.053	5.993	3.915	1.057
Deposits	Log of customer deposits	0.000	12.832	4.279	2.547

Loans	Logarithm of loans	1.281	12.975	5.766	2.431
Assets	Logarithm of assets	0.000	13.730	6.965	1.996
Dummy variables equalling 1 and zero otherwise					
Sav	Savings banks	0	1	0.256	0.436
Com	Commercial banks	0	1	0.341	0.474
Coop	Cooperative banks	0	1	0.401	0.490
BE	Banks operating in Belgium	0.000	1.000	0.0253	0.1572
DK	Banks operating in Denmark	0.000	1.000	0.6163	0.2405
FN	Banks operating in Finland	0.000	1.000	0.00633	0.7937
FR	Banks operating in France	0.000	1.000	0.19182	0.3938
GE	Banks operating in Germany	0.000	1.000	0.1791	0.3835
GR	Banks operating in Greece	0.000	1.000	0.0126	0.1118
IR	Banks operating in Ireland	0.000	1.000	0.02822	0.1656
LU	Banks operating in Luxembourg	0.000	1.000	0.10483	0.3064
NL	Banks operating in the Netherlands	0.000	1.000	0.3398	0.1812
PT	Banks operating in Portugal	0.000	1.000	0.2016	0.1405
SP	Banks operating in Spain	0.000	1.000	0.8467	0.2784
SW	Banks operating in Sweden	0.000	1.000	0.0092	0.0955
UK	Banks operating in the UK	0.000	1.000	0.1054	0.3071

We estimate a stochastic translog cost function with input prices (P), output descriptors (Y) and a trend (t).

$$\log C_{it} = \tau_0 + \tau_1 t + \frac{1}{2}\tau_2 t^2 + \sum_{k=1}^{m} \alpha_k \log Y_{kit} + \sum_{j=1}^{n} \beta_j \log P_{jit} + \frac{1}{2}\left[\sum_{k=1}^{m}\sum_{r=1}^{m} \pi_{kr} \log Y_{kit} \log Y_{rit} + \sum_{j=1}^{n}\sum_{s=1}^{n} \delta_{js} \log P_{jit} \log P_{snt}\right] + \sum_{k=1}^{m}\sum_{j=1}^{n} \theta_{kj} \log Y_{kit} \log P_{jit} + \eta Type_{it} + \kappa Country_{it} + (V_{it} + U_{it}),$$

where C_{nt} is the natural logarithm of variable costs; t is a trend; $\log Y_{int}$ is the natural logarithm of the i-th outputs (deposits, loans, assets) from bank n in period t; $\log P_{jnt}$ is the natural logarithm of the jth input price (wages, capital) from bank n-th in period t. A "Type" dummy defining the type of bank (savings, commercial and cooperative) and a "Country" dummy are also included. $\tau_0, \tau_1, \tau_2, \alpha_k, \beta_j, \pi_{kr}, \delta_{js}, \theta_{kj}, , \eta, \kappa$ are the coefficients to be estimated. The adopted specification is the cost frontier model, known as the error components model in Coelli, Rao and Battese (1998). Table 5 presents the results obtained for the stochastic production frontier, using a GAUSS routine.

The Translog equation was estimated imposing symmetry and linear homogeneity in prices (Cornes, 1992). The latter requires dividing monetary values by the input price of capital-premises (Cornes, 1992), which corresponds to the following restrictions:

$$\sum_{j=1}^{n} \beta_j = 1; \sum_{j=1}^{n} \delta_{js} = 0 \text{ for all } s; \quad \sum_{j=1}^{n} \theta_{kj} = 0, \text{ for all k;} \quad \text{and} \quad \sum_{j=1}^{n} \rho_j = 0.$$

Young's theorem requires the following symmetry restrictions:

$$\pi_{kr} = \pi_{rk} \text{ for all k and r,} \quad \text{and} \quad \delta_{js} = \delta_{sj} \text{ for all j and s.}$$

These restrictions reduce the number of parameters to be estimated. Moreover, the cost function must be non-increasing and convex with respect to the level of fixed input, and non-decreasing and concave with respect to prices of the variable inputs. These conditions were not imposed, but they can be tested to determine whether the cost function is well-behaved at each point within a given data set.

To allow direct interpretation of the first order Translog parameters as elasticities evaluated at the sample mean, every series was divided by its average value (Coelli et al. 1998, p. 33). On the basis of the number of observations and exogenous variables, we have chosen the Translog model with a half-normal distribution, which is statistically supported by the data. The error components model is then adopted as suggested by Coelli et al. (1998).

Table 5: Latent Translog panel cost frontier (dependent variable: log Cost)

Variables	Latent class 1	Latent class 2	Latent class 2
Non-random parameters	Coefficients (t-ratio)	Coefficients (t-ratio)	Coefficients (t-ratio)
Constant	0.214 (1.521)	0.351 (3.218)*	1.038 (3.591)*
Trend	0.255 (3.126)*	0.314 (4.217)*	0.419 (3.567)*
Trend2	-0.052 (-3.218)*	-0.051 (-4.216)*	-0.048 (3.038)*
Log PL	0.073 (3.128)**	0.125 (3.673)*	0.129 (3.214)*
LogPK1	0.321 (2.788)*	0.318 (3.782)*	0.325 (3.035)*
Log Deposits	0.521 (3.627)*	0.451 (3.752)*	0.402 (3.318)*

Log Loans	0.487 (3.523)*	0.528 (4.128)*	0.507 (4.072)*
Log Assets	0.215 (3.127)*	0.207 (4.129)*	0.225 (3.521)*
1/2Trend^2	0.218 (4.234)*	0.251 (3.218)*	0.262 (4.519)*
1/2Log PL^2	0.052 (4.215)*	0.065 (4.032)*	0.075 (4.273)*
1/2Log K1^2	0.127 (3.783)*	0.153 (2.832)*	0.187 (3.282)*
1/2 log Deposits^2	0.521 (1.945)	0.637 (3.218)*	0.574 (0.378)
1/2 log loans^2	0.832 (4.278)*	0.763 (3.219)*	0.915 (3.021)*
1/2 log Assets^2	0.832 (3.219)*	0.917 (3.178)*	1.021 (2.917)*
Log PL* Log PK1	0.415 (3.218)*	0.521 (4.128)	0.718 (3.016)
Log PL* Log Deposits	0.518 (3.812)*	0.485 (2.184)**	0.632 (3.218)
Log PL *Log Loans	0.127 (2.583)	0.145 (2.832)	0.183 (3.015)
Log PL *Log Assets	-1.021 (-0.128)	-0.893 (-1.037)	-0.905 (-0.896)
Log PK1* Log Deposits	-0.075 (-1.056)	-0.083 (-2.153)	-0.091 (-3.017)*
Log PK1* Log Loans	0.208 (1.344)	1.551 (2.816)	3.526 (1.231)
Log PK1 * Log Assets	0.454 (7.036)	0.369 (5.675)	0.421 (3.781)*
Log Deposits * Log Loans	0.432 (3.887)	0.459 (2.838)	0.0321 (2.219)**
Log deposits * log Assets	0.058 (2.448)	0.110 (3.054)	0.127 (4.381)*
Log Loans * Log Assets	0.469 (3.488)*	0.214 (2.950)	0.314 (4.214)*
Sav	0.448 (3.260)	-0.368 (-3.021)	-0.416 (-4.783)*
Com	-0.488 (-3.781)	0.511 (3.966)	-0.521 (-4.232)*
Coop	-0.1362 (-4.032)	-0.1225 (-3.079)	0.314 (4.521)*
BE	0.081 (3.675)*	0.437 (5.260)*	0.225 (4.367)*
DK	0.057 (3.791)*	0.032 (0.321)	0.128 (0.762)
FN	0.072 (5.321)*	0.142 (4.403)*	0.021 (4.218)*
FR	0.0321 (3.821)*	0.142 (3.289)*	0.073 (4.452)*
GE	-0.0217 (-3.783)*	0.942 (6.574)*	0.073 (4.452)*
GR	0.172 (5.321)*	0.142 (4.403)*	-0.0452 (-1.295)
IR	-0.0321	0.142	0.0375 (1.219)

	(-1.821)	(3.289)*	
LU	-0.132 (1.045)	0.132 (1.032)	0.135 (3.285)*
NL	0.0217 (1.783)	0.942 (6.574)*	0.088 (3.563)*
PT	0.630 (1.517)	0.282 (5.970)*	-0.325 (-4.378)*
SP	0.135 (1.741)	0.273 (4.233)`*	0.218 (4.893)*
SW	-0.098 (-3.255)*	0.126 (3.673)*	0.218 (3.174)*
UK	0.031 (3.156)*	0.132 (3.125)*	0.052 (1.015)
$\lambda = \sigma_U / \sigma_V$	0.127 (3.218)*	0.102 (3.215)*	0.091 (4.145)*
Log likelihood	1252.132	—	—
Nobs	1554	—	—

(t-statistics) in parentheses are below the parameters. Those followed by * are significant at 1% level. Those followed by ** are significant at 5% level.

The estimated parameters of the latent frontier model and their t-statistics are reported in Table 5. The log-likelihood value of the estimated latent mixed logit model is 1252.132. The overall fit of the model is reasonably good, the Chi-square statistic being equal to 205.123 with 10 degrees of freedom and a significance level of 0.00052.

To summarise the results, it appears that there are three segments in the sample, which are statistically significant. The first one is the more representative, since the probability of a bank belonging to this segment is 0.527, whereas the probabilities for the second and third segment are 0.317 and 0.156 respectively. The segments are all positively related to output and prices in the cost frontier (Varian, 1987). Moreover the trend is positive but grows at a decreasing rate in all cases. The first segment is characterised by a positive relationship with saving banks and a negative relationship with commercial and cooperative banks. The second segment exhibits a positive relationship with commercial banks and a negative one with the other two types of banks. The third cluster can be identified as cooperative banks. As for the country factor, this is mostly positively related to the clusters, with few cases of a negative relationship.

The above findings indicate that latent frontier models describe the European banking system fairly well, when allowing for heterogeneity and defining segments in the sample. This is an important result, since it implies that a common banking policy for all European banks is inappropriate, given the heterogeneity revealed by the three segments. Banking policies should instead be tailored by segments. Note that, given the number of available observations, the model cannot identify more than three segments, but more could in fact exist, with additional heterogeneity. Banks in the first segment can be identified as savings banks, consistently with traditional homogenous frontier models (Williams, Peypoch and Barros, 2007). The second segment corresponds to commercial banks, and all the statistically significant parameters have the same signs as in the first segment. At a European level, commercial banks face intense competition, and these results are consistent with previous research (Peypoch, Barros and Wiliams, 2007). The third segment includes co-operative banks (Molyneux and Williams, 2005).

Overall, our findings are quite intuitive, as European banks are clearly not homogenous. The most homogenous characteristic is the type of bank (commercial, savings and cooperative). This characteristic translates into clusters, and leads to different performance levels. As for the hypotheses of interest, we do not reject any of them. In the case of the first one, this suggests that savings banks are efficient, as indicated by the negative sign (Barney, 1991; Rumelt, 1991), and also found in other studies using other approaches (Williams, Barros and Peypoch, 2007). Similarly, both savings and cooperative banks are found to perform efficiently (see Williamson 1979, 1985, and Caves and Porter, 1977, respectively).

Therefore, banks appear to be efficient in using the unique assets they own and control (Teece et al., 1997). The main factor behind the segmentation is the specific property resource, that creates competition for banks in different segments**.** Strategic-group theory (Caves and Porter, 1977), which accounts for different efficiency scores in terms of differences in the structural characteristics of units within an industry, could also partly explain the efficiency differences observed in the European banking sector.

6. Conclusions

This paper has proposed an econometric framework for the comparative evaluation of European banks and their operational activities which allows for heterogeneity. The estimated latent frontier model appears to be able to capture the dynamics of the data better than standard methods. Significant heterogeneity is confirmed to be present, implying that banking policies designed for specific clusters are more effective than common ones. It is also found that efficiency is a property of all the different types of commercial banks.

The main limitation of our analysis is the fact that the data span is relatively short, restricting the estimation of latent classes to three only: a longer data span would allow the identification of more latent variables. This is left for future research. However, the present study already offers convincing evidence of the segmentation of the European banking sector, and of the resulting need to define business strategies and policies specific to each segment.

References

Alam, I. M. S. (2001). A Non-Parametric Approach for Assessing Productivity Dynamics of Large Banks. Journal of Money, Credit and Banking, 33, 121 – 139.

Aly, Hassan Y.; Grabowski, Richard; Pasurka, Carl; Rangan, Nanda. (1990) Technical, Scale, And Allocative Efficiencies in U.S. Banking: An Empirical Investigation," *Review of Economics and Statistics,* 72, 2, 211-18.

Banker, R.D. and Maindiratta, A. (1988). Nonparametric analysis of technical and allocative efficiency in production. *Econometrica*, 56, 1315-1332.

Barros, C.P.; Peypoch, N. and Williams, J. (2007) A Note on Productivity Change in European Co-operative Banks: The Luenberger Indicator Approach. Mimeo.

Barros, C.P.; Ferreira, C. and Williams, J. (2007) Analysing the Determinants of Performance of Best and worst European banks: A Mixed Logit approach. Journal of Banking and Finance (forthcoming).

Bauer, P.W., Berger, A.N., Humphrey, D.B (1993). Efficiency and Productivity Growth in US Banking, in Fried, H.O., Lovell, K.A.L., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency: Techniques and Applications. Oxford University Press, Oxford, pp. 386-413.

Berger, Allen N.; Humphrey, David B (1997). Efficiency of Financial Institutions: International Survey and Directions for Future Research, *European Journal of Operational Research*, 98, 2, 175-212.

Berger, A.N., Mester, L.J., 2003. Explaining the Dramatic Changes in Performance of US Banks: Technological Change, Deregulation, and Dynamic Changes in Competition. Journal of Financial Intermediation 12, 1, 57-95.

Bos, J.W. B. and Schmiedel, H. (2007) Is there a single frontier in European banking market? Journal of Banking and Finance, 31, 2081-2102

Boussemart, J. P., Briec, W., Kerstens, K., & Poutineau, J. C. (2003). Luenberger and Malmquist productivity indices: Theoretical comparisons and empirical illustration. *Bulletin of Economic Research*, 55, 391-405.

Briec, W. (1997). A graph type extension of Farrell technical efficiency measure. *Journal of Productivity Analysis*, 8, 95-110.

Briec, W., Dervaux, B., Leleu, H., 2003. Aggregation of directional distance functions and industrial efficiency. Journal of Economics 79, 237-261.

Casu, B., C. Giradone and P. Molyneux, (2004) 'Productivity in European Banking – A Comparison of Parametric and Non-Parametric Approaches', Journal of Banking & Finance, 28, 10, 2521-2540.

Casu, B., and P. Molyneux, (2004) A Comparative Study of Efficiency in European Banking," *Applied Economics,* 35, 17, November 20, 1865-76.

Chambers, R.G. (1996). A New Look at Exact Input, Output, and Productivity Measurement. Department of Agricultural and Resource Economics, Working Paper 96-05.

Chambers, R.G. (2002). Exact Nonradial Input, Output, and Productivity Measurement. *Economic Theory*, 20, pp. 751–65.

Chambers, R.G., Chung, Y. and Färe, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, 70, pp. 407-419.

Chambers, R.G., Chung, Y., and Färe, R. (1998). Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications*, 98, pp. 351-364.

Chen, Tser Y.; Yeh, Tsai L. (2000) A Measurement of Bank Efficiency, Ownership and Productivity Changes in Taiwan, *Service Industries Journal*. 20, 1, January 2000, pp. 95-109.

Dietsch, M.; Weill, L**.** (2000) *The Evolution of Cost and Profit Efficiency in the European Banking Industry*, Research in Banking and Finance, JAI Press, 2000.

Diewert, W. (1998). Index Number Theory Using Differences Rather Than Ratios. Department of Economics Discussion Paper 98–10, Vancouver: University of British Columbia.

Diewert, W. (2000). Productivity Measurement Using Differences Rather Than Ratios: A Note. School of Economics Discussion Paper 2000/1, Sydney: University of New South Wales.

Drake, Leigh. (2001) Efficiency and Productivity Change in UK Banking, *Applied Financial Economics,* 11, 5, October 2001, pp. 557-71.

Ericson, R., Pakes, A., 1995. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. Review of Economic Studies 62, pp. 53-82.

Färe, R., Grosskopf, S., Lindgren, B. and Roos, P. (1989). Productivity developments in Swedish hospitals: a Malmquist output index approach, in A. Charnes, V. W. Cooper, A. Lewin and L. Seiford (eds.), Data Envelopment Analysis: Theory, Methodology and Applications, Boston: Kluwer Academic Publishers.

Färe, R. and Grosskopf, S. (2004), *New Directions: Efficiency and Productivity*, Kluwer Academic Publishers, Boston/London/Dordrecht.

Färe, R. and Primont, D. (2003), Luenberger productivity indicators: Aggregation across firms. *Journal of Productivity Analysis*, 20, 425–435.

Farrell, M.J., (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, Series A, 120 (3), 253-290.

Garden, K. A.; Ralston, D. E. (1999) The X-efficiency and Allocative Efficiency Effects of Credit Union Mergers," *Journal of International Financial Markets, Institutions and Money,* 9, 3, 285-301.

Humphrey, D.B., Pulley, L.B., (1997). Banks' Responses to Deregulation: Profits, Technology, and Efficiency. Journal of Money, Credit and Banking 29, 73 –93.

Jensen, M.C., Meckling, W. (1976). Theory of the firm: managerial behaviour, agency costs and capital structure. *Journal of Financial Economics* 3, 305-60.

Jovanovic, B., 1982. Selection and the Evolution of Industry. Econometrica 50, pp. 649-670.

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

Libenstein, H., (1966). Allocative efficiency vs. 'X-efficiency`. *American Economic Review* 56 (3), 392-414.

Luenberger, D.G. (1992). Benefit function and duality. *Journal of Mathematical Economics*, 21, 461-481.

Malmquist, S., (1953) Index Numbers and Indifference Surfaces, *Trabajos de Estatistica*, 4, pp. 209-42.

Mendes, V., Rebelo, J., The Effects of Banks' M&A on Efficiency: The Portuguese Experience. In Margarida Abreu and Vítor Mendes (Eds.), *What Financial System for the Year 2000?* Lisbon: Bolsa de Valores de Lisboa, 2000, pp. 277-89.

Molyneux, P., Altunbas, Y., Gardener, E.P.M. and Moore, B. (2001) Efficiency in European Banking. European Economic Review, 45, 1931-1955.

Molyneux, P. and Williams, J. (2005), 'The productivity of European co-operative banks', *Managerial Finance*, 31, 11, 26-35.

Nelson, R.R., Winter, S.G., 1978. Forces Generating and Limiting Concentration under Schumpeterian Competition. Bell Journal of Economics 9, 524-548.

Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Cambridge, Mass: The Belknap Press of Harvard University Press.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Orea, L. and Kumbhakar, S. (2004) Efficiency measurement using stochastic Frontier latent class model. Empirical Economics, 29, 169-183.

Park, K.H. and Weber, W.L. (2006) a note on efficiency and productivity growth in Korean banking, 1992-2002. Journal of Banking and Finance, 30, 2371-2386.

Peypoch, N.; Barros, C.P. and Williams, J. (2007) A note on productivity change in European commercial banking. Mimeo.

Porter, M. (2000) Location, Competition, and Economic Development: Local clusters in a global economy. Economic Development Quarterly, 14, 1:15-34.

Shephard, R.W. (1970). Theory of Cost and Production Functions. Princeton University Press, Princeton NJ.

Stiroh, K.J., 2000. How did Bank Holding Companies Prosper in the 1990s? Journal of Banking and Finance 24, 1703 –1745.

Valverde, S-C. ; Humphrey, D.B. and Paso, R.L. (2007) Opening the black box: Finding the sources of cost inefficiency. Journal of Productivity Analysis, 27,3, 209-220

Varian, H.R. (1984). The nonparametric approach to production analysis. *Econometrica*, 52, 579-597.

Williams, J. (2001). Financial Deregulation and Productivity Change in European Banking. Revue Bancaire et Financiere 8, 470 – 477.

Williams, J. ; Peypoch, N. and Barros, C.P. (2007) The Luenberger indicator and productivity growth: A note on the European savings banks sector. Mimeo

Worthington, Andrew C.; Dollery, Brian E. (2002) Incorporating Contextual Information in Public Sector Efficiency Analyses: A Comparative Study of NSW Local Governments, *Applied Economics,* 34, 4, 453-64.