Determinants of bank efficiency: Evidence from a semi-parametric methodology

Delis, Manthos D and Papanikolaou, Nikolaos I

January 2009

Determinants of bank efficiency: Evidence from a semi-parametric methodology

Manthos D. Delis [a, b, *],
Nikolaos I. Papanikolaou [a]

[a] *Athens University of Economics and Business, 76 Patission Street, Athens 10434, Greece*
[b] *University of Central Greece, Lambrou Katsoni Square, Levadia 32100, Greece*

Abstract
In this paper, we use a semi-parametric two-stage model to examine the effect of bank-specific, industry-specific and macroeconomic determinants of bank efficiency. This method, proposed by Simar and Wilson (2007), relaxes several deficiencies of previous two-stage analyses, which regress non-parametric estimates of bank efficiency on exogenous determinants. In particular, we propose a bootstrap procedure to be used in the second stage and we compare the results obtained to the equivalents of a Tobit model. We suggest that the Tobit regressions inaccurately provide insignificant estimates for the effect of bank size, industry concentration and economic investment on bank efficiency, a fact that illustrates the power of the new method. Since the effect of these determinants has been ambiguous in previous literature, this may be a desideratum for future research.

JEL classification: C14; G21; P34
Keywords: Bank efficiency; semi-parametric models

Acknowledgements: Helpful research assistance by Sophocles Brissimis, Stephanie Milioti and Mike Tsionas is gratefully acknowledged.

[*] Corresponding author. Tel.: (+30) 210 8203354; fax: (+30) 210 9615868.
Email addresses: delis@aueb.gr (M.D. Delis), nickpap@aueb.gr (N.I. Papanikolaou).

1. Introduction

It has been established that banks, in their role as financial intermediaries, contribute significantly to economic activity in a number of ways. During the last two decades the banking sector has experienced major transformations worldwide in its operating environment. Both external and domestic factors have affected its structure, efficiency and performance. An efficient banking sector is better able to withstand negative shocks and contribute to the stability of the financial system. Therefore, it comes as no surprise that since the publication of the seminal papers by Aigner et al. (1977) and Charnes et al. (1978), both econometric (parametric) and linear programming-based (non-parametric) methods have been employed in the estimation of bank efficiency.

Although econometric models can be estimated either by sampling theory or Bayesian techniques, efficiency measurement in these models rely on the choice of functional forms, distributional assumptions, fixity of parameters, etc. Some of these assumptions are particularly strong in the sense that empirical results are often quite sensitive in them. Albeit the linear programming methods do not make such assumptions, they do not separate inefficiency from noise, and this leads to two major shortcomings as regards their usefulness. First, applied researchers are unable to make statements regarding the statistical properties of the estimated functions, while the efficiency scores derived may not be easily analyzed into their various determinants. The second shortcoming seems to be of particular importance if the objective is to examine various hypotheses concerning exogenous determinants of bank efficiency. Many studies (not necessarily restricted to the banking industry) overlook this problem and employ a two-stage approach, where efficiency is estimated in the first stage, and then the estimated scores are regressed on covariates (by specifying censored models), typically different from those used in the first stage. However, in an important contribution, Simar and Wilson (2007), SW hereafter, suggest that these second-stage results are invalid, mainly because the efficiency scores derived in the first stage are serially correlated. In fact, SW suggest a bootstrapping procedure (to be used instead of the censored regression) that permits valid inference and improves on statistical efficiency.

In the present paper, we draw on the technique of SW to provide a robust semi-parametric model of the analysis of bank efficiency into a number of determinants. In particular, efficiency is derived via the most commonly used non-parametric method, i.e. Data Envelopment Analysis (DEA), and then the scores obtained are linked to bank-specific, industry-specific and macroeconomic determinants of bank efficiency, using an advanced bootstrapping technique. We opt for an application of this empirical model to 10 newly acceded EU countries, since their transition from centrally planned to market economies involved quite uniform but immense institutional, structural and managerial changes in their banking sectors. The twelve years of data used (1994-2005) capture almost the entire course of the banking-sector reform process of the countries examined. Finally, the results obtained are compared to the equivalents of censored (Tobit) regressions, in an effort to relate this study with econometric theory and previous findings.

The rest of this article proceeds as follows. Section 2 presents the two-stage methodology employed in the paper. Section 3 describes the banking sectors of the newly acceded EU countries and the determinants of bank performance. Section 4 discusses the empirical results and, finally, Section 5 concludes the paper.

2. Methodology

At least four methods have been put forth to describe factors that could influence the efficiency of a firm, where such factors are not traditional inputs of production and are beyond the influential control of the manager (see Coelli et al., 2005). The most favored of these methods involves a two-stage semi-parametric process.[1] In the first stage, a set of observed pairs of inputs x_i and outputs y_i is used to derive efficiency scores θ_i for all banks $i = 1,\ldots,n$. In the second stage, the efficiency estimates obtained are regressed on exogenous (or even in some cases on endogenous) determinants using censored (Tobit) regression. Yet, as Simar and Wilson (2007) point out this approach presents at least two problems for inference. First, θs are serially correlated in a complicated and unknown way, since they depend on the inputs and outputs of the first-stage analysis (what is used to estimate θ) and also depend on the error term of the second-stage regression. Therefore, the error term depends on the first stage inputs and outputs of the production process. Second, this means that the error term of the censored regression is also correlated with the exogenous determinants. The fact that maximum likelihood is used in the second stage implies that this correlation disappears asymptotically (so that the estimates will be consistent), yet at a very slow rate (which yields invalid inference).

In this section we describe a better methodology of estimation and inference in two-stage semi-parametric models of production processes, building on SW to relate their model to the banking industry. In the first stage, we employ simple input-oriented DEA[2]

[1] The rest of the methods implied are subject to more severe criticism than the one that involves a censored model (see Coelli et al.,2005; Simar and Wison, 2007).

[2] DEA may be computed either as input or output oriented. Input-oriented DEA shows by how much input quantities can be reduced without varying the output quantities produced. Output-oriented DEA assesses by

to measure variable returns to scale productive efficiency of banks. In Stage 2 we describe a bootstrap procedure that accounts, inter alia, for the serial correlation of the first-stage efficiency estimates, thus improving on statistical efficiency. Note that efficiency scores are derived on a country-specific basis so as to avoid incorporating the effect of the different economic environments of our sample into the estimated scores.

Stage 1

To introduce some notation, let us assume that for *N* banks there exist *M* inputs producing *S* outputs. Hence, each bank *n* uses a nonnegative vector of inputs denoted $x^n = (x_1^n, x_2^n, ..., x_m^n) \in R_+^M$ to produce a nonnegative vector of outputs, denoted $y^n = (y_1^n, y_2^n, ..., y_S^n) \in R_+^S$. Production technology, $F = \{(y, x) : x \text{ can produce y}\}$, describes the set of feasible input-output vectors, and the input sets of production technology, $L(y) = \{x : (y, x) \in F\}$ describe the sets of input vectors that are feasible for each output vector (Kumbhakar and Lovell, 2000).

To measure variable returns to scale productive efficiency we use the following input-oriented DEA model, where the inputs are minimized and the outputs are held at their current levels:

how much output quantities can be proportionally increased without changing the input quantities used. The two measures provide the same results under constant returns to scale but give slightly different values under variable returns to scale. Nevertheless, both output- and input-oriented models will identify the same set of efficient/inefficient banks.

$$
\begin{aligned}
&\theta^* = \min \theta, \text{ s.t.} \\
&\sum_{j=1}^{n} \lambda_j x_{ji} \le \theta x_{i0} \quad i = 1,2,\ldots,m; \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{r0} \quad r = 1,2,\ldots,s; \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \qquad j = 1,2,\ldots,n;
\end{aligned} \tag{1}
$$

where $bank_0$ represents one of the N banks under evaluation, and x_{i0} and y_{r0} are the ith input and rth output for $bank_0$, respectively. If $\theta^* = 1$, then the current input levels cannot be proportionally improved, indicating that $bank_0$ is on the frontier. Otherwise, if $\theta^* < 1$, then $bank_0$ represents an inefficient bank and θ^* represents its input-oriented efficiency score. Finally, λ is the activity vector denoting the intensity levels at which the S observations are conducted. Note that this approach, through the convexity constraint $\vec{1}\lambda = 1$ (which accounts for variable returns to scale) forms a convex hull of intersecting planes, since the frontier production plane is defined by combining some actual production planes.

Stage 2

The idea is to estimate a function of the form

$$\theta_i = \alpha z_i + \varepsilon_i \tag{2}$$

where θ are bank efficiency scores derived above, z are a vector of environmental variables that explain variations in efficiency and α are parameters to be estimated. Estimation of Eq. (2) is completed in the three following steps, derived from SW.

1. Obtain maximum likelihood estimates $\hat{\alpha}_k$ of α_k and $\hat{\sigma}_u$ of σ_u in the endogenous truncated regression of θ on its k determinants (z_i) in Eq. (2), where θ<1.
2. Loop over the next three steps L=2000 times to obtain a set of bootstrap estimates $\mathrm{B}_i = \left[(\hat{\alpha}^*, \hat{\sigma}_u^*)_b\right]_{b=1}^{L}$:

   2.1 For each bank i=1,…,m, draw u_i from the $N(0,\hat{\sigma}_u^2)$ distribution with left-truncation at $(1-z_i\hat{a})$. For details on how to draw from a left-truncated normal distribution see the Appendix of Simar and Wilson (2007).

   2.2 Again for each i=1,…,m, compute $\theta_i^* = z_i\hat{\alpha} + u_i$.

   2.3 Use the maximum likelihood method to estimate the endogenous truncated regression of θ_i^* on z_i, yielding estimates μ_μ^*, v_v^*.
3. Use the bootstrap values in B and the original estimates α, σ_u to construct estimated confidence intervals for each element of α and for σ_u. This is done by using the jth element of each bootstrap value $\hat{\alpha}^*$ to find values μ_π^*, v_π^* such that $\Pr\left[-v_\pi \le (\hat{\alpha}_j^* - \hat{\alpha}_j) \le \mu_\pi^*\right] \approx 1-\pi$, for some small conventional value of π, $\pi = 0.05$ in the present analysis. The approximation improves as $L \to \infty$. Substituting μ_π^*, v_π^* for μ_π, v_π in $\Pr\left[-v_\pi \le (\hat{\alpha}_j - \alpha_j) \le \mu_\pi\right] = 1-\pi$ leads to an estimated confidence interval $(\hat{\alpha}_j + \mu_\pi^*, \hat{\alpha}_j + v_\pi^*)$.

The first maximum likelihood estimation will give consistent estimates of α, though without the usual convergence rate. The second maximum likelihood estimation in step 2 is a parametric bootstrap of a nonlinear regression model, whose properties have been thoroughly examined in the past (e.g. Wu, 1986). Therefore, the whole procedure assures consistency and greatly improves on inference.

Estimation of this algorithm is relatively easy with existing software that have embedded commands for truncated regression models. The results derived from this procedure will be compared with the results of a Tobit regression on Eq. (2), in an effort to make statements about the robustness of our procedure, given economic and statistical theory. In what follows, we discuss the specifics of the banking sectors used in the empirical analysis and we present the potential determinants of bank efficiency, along with expectations regarding their effect.

3. Economic environment and data

3.1. Background

Banking system restructuring was quite profound over the last decade in the newly acceded EU countries. Since the mid 1990s their banking systems were extensively reformed through the abolition of administrative interventions and regulations, which seriously hampered its development. The reforms were adopted gradually and supported the further improvement of the institutional framework and the more competent functioning of banks and financial markets in general. The objective of these countries' participation in the EU initiated efforts towards the further deregulation of their banking systems and macroeconomic convergence. During the past few years, banks tried to strengthen their position in the domestic market and acquire a size, partly through M&As, that would allow them to exploit economies of scale, and have easier access to international financial markets.

Banks operating in the countries examined are gradually reaching the standards of their counterparts in the rest of the EU countries. The institutional reforms briefly described above have been viewed as a means to reduce bank costs, particularly those associated with risk management and the evaluation of credit information. However, for smaller and private domestic banks, risk management techniques need to improve further (see EBRD, 2006). In fact, lending in emerging markets is greatly influenced by how banks perceive the legal environment, and the level of hedging against risks that this environment provides. Institutional improvements, such as effective systems for taking collateral and repossessing assets in cases of default, will play a fundamental role in the further development and stability of the banking sector. However, given the restructuring that took place in the last decade, the newly acceded EU countries provide an excellent case for the study of the determinants of bank efficiency.[3]

3.2. Dataset

The model presented in Section 2 is estimated on a panel of banks from 10 newly acceded EU countries (listed in Table 1), which corresponds to a relatively long period that covers the banking sector reform process in these countries (namely the 1994-2005 period). We choose to limit the empirical analysis to the unconsolidated statements of commercial, savings and cooperative banks in order to reduce the possibility of introducing aggregation bias in the results. All bank-level data used are obtained from the BankScope database. During the sample period a number of M&As and bank failures took place, which are taken into account in our dataset so as to avoid selectivity bias.

[3] For a detailed review of the reform process in the CEE countries' financial sectors see various issues of the EBRD Transition reports (e.g. Transition report 2006: Finance in transition).

Also, the data were reviewed for reporting errors or other inconsistencies (zero or negative values for the variables used). This yielded an unbalanced panel dataset of 4368 observations corresponding to 364 banks. All bank-level data are reported in euros and are expressed in constant 1994 prices (using individual country GDP deflators). Below we discuss the variables used to estimate Program (1).

The first problem encountered in evaluating bank efficiency is the definition and measurement of bank output. The two most widely used approaches are the 'production' and the 'intermediation' approaches.[4] While we acknowledge that it would probably be best to employ both approaches to identify whether the results are biased when using a different set of outputs, sufficient data to perform such an analysis on banks from newly acceded EU countries is generally unavailable. Hence, this study uses the 'intermediation approach' for two main reasons: First, this approach is inclusive of interest expenses that usually account for over one-half of total costs, and second the BankScope database lacks the necessary data for implementing the production approach. Accordingly, this study specifies two outputs, namely total loans and total securities; and two inputs, i.e. operating expenses (non-interest and personnel expenses) and total deposits and short-term funding.[5] Both inputs and outputs have risen considerably during the sample period

[4] Under the former approach output is measured by the number of transactions or documents processed over a given time period (see Berger and Humphrey, 1997). Under the latter approach output is measured in terms of values of stock variables (such as loans, deposits, etc.) appearing in bank accounts.

[5] The definition of inputs and outputs varies widely across studies of bank efficiency. In this paper, given the limitations of the BankScope database, the further disaggregation of inputs and outputs is not possible (i.e. personnel expenses or fixed assets are not reported for many banks). Clearly it is possible that the use of expenses rather than physical inputs could result in some bias against those banks that employ high quality and therefore high cost inputs. This potential bias should be mitigated, however, given that banks with high quality inputs should expect to see some benefit in output terms. Hence, if high quality inputs are sufficiently productive, such banks will not be disadvantaged from a relative efficiency perspective (see Berger and Humphrey, 1997; Drake and Hall, 2003). Also, some studies suggest that deposits have both input and output characteristics (e.g. Berger and Humphrey, 1997). However, even this separation of deposits is difficult, given the diversity of the banking systems examined. For the shake of comparison, total deposits are treated here as inputs.

due to M&As and the quickly growing size of banking institutions (especially of the newly established foreign institutions) of the region (summary statistics are presented in Table 1).

3.3. Determinants of efficiency

In the literature, bank efficiency is usually expressed as a function of internal and external determinants. The internal determinants originate from bank accounts (balance sheets and/or profit and loss accounts) and therefore could be termed micro or bank-specific determinants of efficiency. The external determinants are variables that are not related to bank management but reflect the economic and legal environment that affects the operation and performance of financial institutions. A number of explanatory variables have been proposed for both categories, according to the nature and purpose of each study.

Studies dealing with internal determinants employ variables such as size, capital and risk management. One of the most important questions underlying bank policy is which size optimizes bank efficiency. Generally, the effect of a growing size on efficiency has been proved to be positive to a certain extent. However, for banks that become extremely large, the effect of size could be negative due to bureaucratic and other reasons. Hence, the size-efficiency relationship may be expected to be non-linear. We use the banks' real assets (logarithm) and their square in order to capture this possible non-linear relationship.

The need for risk management in the banking sector is inherent in the nature of the banking business. Changes in credit risk may reflect changes in the health of a bank's loan portfolio (see Cooper et al., 2003), which may affect the performance of the institution, since poor asset quality is the single most important cause of bank failures. During periods of increased uncertainty, financial institutions may decide to diversify their portfolios in order to reduce their risk. However, the results of the existing literature are better described as mixed, with studies like Altunbas et al. (2000) suggesting that efficiency is not very sensitive to credit risk, and others like Hughes and Mester (1993) reporting an opposite result.[6] To proxy credit risk we use the loan-loss provisions to loans ratio (PL).

Turning to the external determinants of bank efficiency, it should be noted that we can further distinguish between control variables that describe the macroeconomic environment, such as inflation, interest rates and gross domestic product, and variables that represent market characteristics. The latter basically refer to banking-sector reform, market concentration and ownership (see Table 1 for summary statistics).

Data on the banking reform process are obtained from the EBRD. In particular, we use the EBRD index of banking sector reform, either as a structural index (*ebrd*) or to generate time dummy variables. This index has been compiled by the EBRD with the primary purpose of assessing the progress of the banking sectors of formerly centrally planned economies. As this indicator quantifies and qualifies the degree of liberalization of the banking industry, it is suitable for an explicit evaluation of the effect of banking sector reform on the performance of banks. Related studies simply measure the impact of

[6] These studies generally use parametric approaches to examine the effect of credit risk, which is treated as an endogenous variable in the production process. Acknowledging that this may be a preferred strategy, we still use it as a control variable in the empirical analysis of the determinants of bank efficiency, mainly because provisioning also has clear implications regarding the stability of the financial environment, within which banks operate (effects that may be clear of managerial control). The extent to which the one effect dominates over the other is beyond the scope of the present paper. The same assumption also carries on to bank size, which is the other bank-specific determinant employed in this paper.

deregulation (or specific deregulation policies) on bank performance or competition; they do not focus on the reform process as a whole.[7] The values of *ebrd* range from 1.0 to 4.0+, with 1.0 indicating a rigid centralized economy and 4.0+ implying the highest level of reform, which corresponds to a fully industrialized market economy. The criteria used for the compilation of the index are common to all countries (see EBRD Transition reports, various issues). When the index is used to formulate dummy variables, we assume that changes in the regulatory regime remain over time, and thus the (country-specific) dummies take a value of one at the year of the change and remain equal to one until the end of the sample period. Obviously, the reform process when treated like this is viewed as an ongoing process that affects banks not only at the year of change in the regulatory regime, but for all the succeeding years of the sample period (see Salas and Saurina, 2003). The upward trend of the index reflects the extensive restructuring that took place in the banking sectors examined during the sample period.

The new regulatory framework in the newly acceded EU countries significantly increased the attractiveness of its banking system for foreign investors. In the period under consideration there was a notable entry of foreign banks, which were looking for acquisition opportunities in the promising banking systems of the countries examined. Foreign ownership may have an impact on bank efficiency due to a number of reasons: First, the capital brought in by foreign investors decrease fiscal costs of banks' restructuring (Tang et al., 2000). Second, foreign banks may bring expertise in risk

[7] For instance, Salas and Saurina (2003) and Kumbhakar and Lozano-Vivas (2005) employ all the deregulation events that occurred in the period under examination to capture the deregulation process in the Spanish banking industry. Angelini and Cetorelli (2003) measure deregulation via changes in minimum capital requirements, or through the abrogation of the interest rate ceilings policy. Similarly, Yildirim and Philippatos (2007) choose foreign bank penetration to capture deregulation. Other studies also use abolition of entry restrictions as deregulatory proxies (e.g. Demirguc-Kunt et al., 2004).

management and a better culture of corporate governance, rendering banks more efficient (Bonin et al., 2005). Third, foreign bank presence increases competition, driving domestic banks to cut costs and improve efficiency (Claessens et al., 2001). Finally, domestic banks have benefited from technological spillovers brought about by their foreign competitors. For these reasons, an examination of the impact of foreign ownership on the efficiency of banks is a useful exercise and this potential effect is captured by the evolution of the shares of foreign banks as a percent of the total bank assets (constructed on a country-specific basis).

A relationship between bank efficiency and ownership may also exist due to spillover effects from the superior performance of privately-owned banks compared with publicly-owned banks, which do not always aim at efficiency. Although there is no clear empirical evidence to support such a view, the peculiarity of the banking sectors examined, where the share of commercial banks under public ownership was relatively high until the early 2000s makes the examination of the hypothesis appealing. To test this hypothesis, we use the time-dependent market share (in terms of assets) of publicly-owned banks in the sector (once again this variable is constructed on a country-specific basis).

The efficient structure paradigm links concentration to high profitability through efficiency (Demesetz, 1973). This hypothesis posits that relative efficient banks compete more aggressively for market shares, which leads to more concentrated markets. Yet, other studies showed that in highly concentrated markets, risk aversion prevails, rendering the relationship between efficiency and concentration negative (see e.g. Sathye, 2001). This possible adverse relationship is further enhanced if the "quiet life" hypothesis holds; that is increased concentration leads to a relaxed banking environment with no incentives to minimize costs (Berger and Humphrey, 1997). In an effort to identify the effect of concentration on bank efficiency, we use a concentration ratio constructed from the market shares (in terms of assets) of the 3 bigger banks in each country.

Finally, following the literature, the second-stage analysis includes some macroeconomic country-specific variables, namely the ratio of total investment to GDP (*invgdp*) as a proxy for fluctuations in economic activity, and a short-term interest rate (*ir*), which captures variability of market interest rates. These variables are taken from the EBRD and the WDI. Bank efficiency is sensitive to macroeconomic conditions despite the trend in the industry towards greater geographic diversification and larger use of financial engineering techniques to manage risk associated with business cycle forecasting. Generally, higher economic growth encourages banks to lend more and permits them to charge higher margins, as well as improving the quality of their assets.

4. Estimation results

Table 2 reports average findings of the non-parametric first stage, whereas Table 3 presents the effect of bank-specific, industry-specific and macroeconomic determinants of bank efficiency. Note that the analysis of Stage 1 is carried out for each country separately and therefore the mean efficiency scores presented in Table 2 only reflect the dispersion of efficiency within each sample; that is, they tell us nothing about the efficiency of one sample relative to another. The results indicate that the banking sectors of almost all the newly acceded EU countries show a gradual improvement in their efficiency levels. This is not surprising, since the banking systems examined have seen fundamental changes in their ownership structure (private vs. public, foreign vs.

domestically owned banks), including mergers. In addition, the relatively stable macroeconomic conditions of the period, coupled with a significant improvement in the operating expenses management, may have also led to improved efficiency. The majority of banks comprising the sample seem to cluster around levels of efficiency of approximately 65%, which is a score similar to that found in other recent nonparametric analyses of western European banking systems (see e.g. Casu and Molyneux, 2003). Among the countries examined the most significant efficiency gains over the examined period are found for Estonia, Hungary and Slovenia. On the other end, Latvia and Poland present lower average efficiency scores in 2005 compared with 1994.

In Table 3 we report estimation results of the Stage 2 analysis. We present three different specifications (denoted by SW), which are also estimated using a simple Tobit model. In the first specification, we provide the results of a basic model that does not include the bank-specific variables. In the second, we augment the model with these variables, while in the third we replace the EBRD index with the dummies that correspond to changes in the index. Coefficients among the different specifications (on a model-specific basis) have fairly stable coefficients, with the SW model usually suggesting higher coefficients than the Tobit model.[8]

Turning to the discussion of the explanatory power of the variables employed, we first note that a clear positive relationship is documented between the EBRD index of banking sector reform and bank efficiency, revealing that the latter has gained ground substantively throughout the reform period. The coefficients on *ebrd* indicate an average percentage increase of about 18.5 basis points for productive efficiency when we use the Tobit model (see columns 1b and 2b), while the increase is 2 basis points higher and

[8] This result is aligned with the discussion in Simar and Wilson (2007).

more statistically significant when the semi-parametric model is employed (see columns 1a and 2a). This finding is verified when the reform binary variables are used instead of the structural indicator. In fact, column (3a) shows a positive and significant effect on bank efficiency of the reform measures undertaken in the years 1997, 2000, 2003, 2004 and 2005, while a non-significant impact is documented for the rest of the reform initiatives. When the Tobit model is employed (column 3b) there are two disparities: the reforms undertaken in 2001 seem to have a negative and significant impact on efficiency, whereas the reforms undertaken in 2004 do not significantly affect it.

Concerning the quality measure of credit risk, the regression results reveal a negative relationship with bank efficiency, confirming our expectations. The use of a Tobit model provides no differentiation to this result, except that once again the Tobit estimates are less statistically significant. Serious financial problems have arisen in the region from the failure of banks to recognize impaired assets and create reserves for writing-off these assets. An immense help towards smoothing these anomalies would be provided by improving the transparency of the financial systems, which in turn will assist banks to evaluate credit risk more effectively, avoid problems associated with hazardous exposure and, therefore, raise their productive efficiency.

A noteworthy result is that the size of the banking institutions plays a significant (economic and statistical) role on their efficiency only if we employ the SW model. In particular, and given the negative effect of squared size variable, the relationship between bank size and efficiency is non-linear, with efficiency increasing with size to a certain point and decreasing thereafter. This finding provides great support to the mainstream view, according to which large banks hire more efficient managers, who succeed in their attempt to establish scale and scope economies. However, this effect is not substantiated by the Tobit regressions (see columns 2b and 3b) and is probably lost in the serial correlation between the DEA estimates. In fact, the simple truncated regression (first step of Stage 2), also suggests an insignificant relationship between size and efficiency, however the effect is strengthened throughout the rest of the procedure.[9]

The empirical results also show that industry concentration has a clearly negative impact on bank efficiency only when the SW model is employed; in contrast, Tobit estimates are both economically and statistically insignificant. The difference between the two models is large with the respective average coefficients being 0.12 and 0.03. Once again, this evidence may suggest that the serial correlation between DEA estimates imposes a serious downward bias in the effect of banking industry concentration (at least in our sample), which is an important element in certain hypotheses, like the structure-conduct-performance and the efficient-structure hypotheses. Undoubtedly, further research is required if one is to make clear statements on the verification of these hypotheses.

The signs of the coefficients on *pub* are negative, indicating that public ownership has an adverse effect on efficiency, which is statistically significant at the 1 % level. In contrast, foreign ownership has a strong positive impact on efficiency.[10] It seems that technological innovations and new managerial practices brought about by foreign-owned institutions have substantially facilitated the efficiency advancement. Our results are in line with those of numerous studies considering the relationship between ownership and performance in transition economies. We indicatively mention that of Grigorian and

[9] The initial results are available upon request.

[10] Note that the effect of ownership is stronger when the Tobit model is used. This suggests that the SW model does not necessarily imply increased significance of the results.

Manole (2002), who employed a DEA method to estimate bank efficiency in seventeen transition economies, and that of Claessens et al. (2001), who found that foreign participation reduces the overhead costs in a set of both developed and developing countries.

As discussed above, we use the investment to GDP ratio and a short-term interest rate to proxy the country-specific macroeconomic and monetary conditions respectively. The results reveal a positive and significant link between efficiency and *ir*, as well as a positive relation with *invgdp*. Yet, the latter result holds only when regressions are run with the SW model. Under Tobit analysis, the coefficients on *invgdp* are not statistically significant. This is the third determinant of efficiency that is identified with enhanced significance when using the SW model, which naturally implies that correction of the statistical problems associated with the Tobit model is particularly important in the robustness of the estimates.

5. Conclusions

Along the lines of previous research we have analyzed the bank efficiency estimates derived from DEA on certain bank-specific, industry-specific and macroeconomic determinants. However, contrary to standard accounts, we introduced the two-stage model of Simar and Wilson (2007) into a large scale empirical analysis of the banking systems of newly acceded EU countries. In the first stage, we employ simple DEA to measure the productive efficiency of banks and in the second stage we analyze the effect of the aforementioned determinants by carrying out an advanced bootstrap procedure. This procedure offers a solution to the many problematic features of censored regressions, the most important being the serial correlation of the first-stage efficiency scores.

The results suggest that the bootstrapping technique unmasks some of the explanatory power of certain variables. In particular, bank size is found to have a positive significant economic and statistical effect on bank efficiency only when the SW model is employed, losing its entire significance when the Tobit model is used instead. The same pattern is documented for the effect of industry concentration (negative effect under the SW model) and the investment environment (positive effect). We contend that these results have important implications for the relevance of well-known hypotheses that refer to the performance of the banking sectors, like the structure-conduct-performance and the efficient structure hypotheses. Since the significant variation in the results of the two models may also have important implications to other determinants of bank efficiency, not considered in the present analysis, this may be a desideratum for future research.

References

Aigner D., Knox Lovell, C.A. and Peter Schmidt P. (1977), "Formulation and Estimation of Stochastic Frontier Production Models", *Journal of Econometrics*, Vol. 6, pp. 21-37.

Altunbas, Y., Liu, Ming-Hau, Molyneux, P. and Seth, R., (2000), "Efficiency and Risk in Japanese Banking", *Journal of Banking and* Finance, Vol. 24, pp. 1605-28.

Angelini, P. and Cetorelli N. (2003), "The effects of regulatory reform on competition in the banking industry", *Journal of Money, Credit, and Banking*, Vol. 35, pp. 663-84.

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

Bonin, J.P., Hasan, I. and Wachtel, P. (2005), "Bank performance, efficiency and ownership in transition countries", *Journal of Banking and Finance*, Vol. 29, pp. 31-53.

Casu, B. and Molyneux, P. (2003), "A Comparative Study of Efficiency in European Banking", *Applied Economics*, Vol. 35, pp. 1865-76.

Charnes, A., Cooper W. and Rhodes E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research*, Vol. 2, pp. 429-44.

Claessens, S., Demirguc-Kunt, A. and Huizinga H. (2001), "How does foreign entry affect domestic banking markets?", *Journal of Banking and Finance*, Vol. 25, pp. 891-911.

Coelli, T., Rao, D.S.P., O'Donnell, C.C. and Battese, G.E. (2005), *An Introduction to Efficiency and Productivity Analysis*, Springer, New York.

Cooper, M., Jackson, W. and Patterson G. (2003), “Evidence of predictability in the cross-section of bank stock returns”, *Journal of Banking and Finance*, Vol. 27, pp. 817-50.

Demirguc-Kunt, A., Laeven, L. and Levine R. (2004), “Regulations, market structure, institutions, and the cost of financial intermediation”, *Journal of Money, Credit, and Banking*, Vol. 36, pp. 593-622.

Demsetz, H. (1973), “Industry structure, market rivalry, and public policy”, *Journal of Law and Economics*, Vol. 16, No. 1. pp. 1-9.

Drake, L. and Hall, M.J.B. (2003), “Efficiency in Japanese banking: An empirical analysis”, *Journal of Banking and Finance*, Vol. 27, pp. 891-917.

EBRD Transition Report 2006, European Bank for Reconstruction and Development, London.

Grigorian, D.A. and Manole, V. (2002), “Determinants of commercial bank performance in transition: an application of Data Envelopment Analysis”, *IMF Working Paper*, WP/02/46.

Hughes, J.P. and Mester, L.J. (1993), “A quality and risk adjusted cost-function for banks: Evidence on the ‘too-big-to-fail’ doctrine”, *Journal of Productivity Analysis*, Vol. 4, pp. 293-315.

Keeley, M.C. (1990), “Deposit insurance, risk, and market power in banking”, *American Economic Review*, Vol. 80, pp. 1183-1200.

Kumbhakar, S.C. and Lovell, C.A.K. (2000), *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge.

Kumbhakar, S.C. and Lozano-Vivas, A. (2005), “Deregulation and productivity: The case of Spanish banks”, *Journal of Regulatory Economics*, Vol. 27, pp. 331-51.

Miller, S. and Noulas A. (1997), “Portfolio mix and large-bank profitability in the USA”, *Applied Economics*, Vol. 29, pp. 505-12.

Salas, V. and Saurina, J. (2003), “Deregulation, market power and risk behavior in Spanish banks”, *European Economic Review*, Vol. 47, pp. 1061-75.

Sathye, M. (2001), “X-efficiency of Australian banking: An empirical investigation”, *Journal of Banking and Finance*, Vol. 25, pp. 613-30.

Simar, L. and Wilson, P.W. (2007), “Estimation and inference in two-stage, semi-parametric models of production processes”, *Journal of Econometrics*, Vol. 136, pp. 31-64.

Tang, H., Zoli, E. and Klytchnikova, I. (2000), “Banking crises in transition economies. Fiscal costs and related issues”, *World Bank Policy Research Working Paper*, No. 2484.

Yildirim, H.S. and Philippatos, G.C. (2007), “Restructuring, consolidation and competition in Latin American banking markets”, *Journal of Banking and Finance* , Vol. 31, pp. 629-39.

Wu, C.F.J. (1986), “Jackknife bootstrap and other resampling methods in regression analysis”, *Annals of Statistics*, Vol. 14, pp. 1261-95.

Table 1
Descriptive statistics by country and through time

Country	Deposits	Operating expenses	Loans	Securities	Assets	Concent.	ebrd	ir	invgdp
Bulgaria	369543	40119	173574	52388	699003	0.658	2.834	26.850	17.492
	1153297	114744	508450	194806	1612874	0.181		50.263	5.747
Czech	826647	64813	245522	471046	2769430	0.736	3.362	7.425	29.150
	2333851	188499	615576	2242569	6740681	0.146		4.951	1.999
Estonia	148847	15193	83598	14177	329704	0.985	3.501	11.239	28.317
	249862	34126	184138	39222	349397	0.011		6.525	1.836
Hungary	261343	20341	150198	41739	661711	0.855	3.667	16.050	22.125
	687891	40828	448350	199562	1229588	0.101		7.634	1.574
Latvia	221715	19882	143328	54074	513691	0.863	3.223	8.583	24.391
	648897	44016	466209	165898	998002	0.055		10.148	6.126
Lithuania	119094	9898	74679	49512	435659	0.671	3.028	8.958	22.392
	317985	17398	252820	184244	653681	0.099		6.644	2.135
Poland	133195	18825	81301	32103	410206	0.550	3.248	15.567	21.267
	696864	46525	368764	273076	1232666	0.028		7.901	2.607
Romania	365849	19960	183360	90259	810592	0.586	2.696	31.908	21.833
	967780	41214	542599	375642	1587698	0.143		12.033	2.787
Slovakia	316425	23273	103444	118123	778997	0.850	3.029	10.850	28.342
	1027710	55965	316492	835073	1812086	0.143		7.118	4.137
Slovenia	389928	31450	352859	71019	1745472	0.742	3.193	9.325	24.392
	1190925	69622	1473242	273222	5335096	0.079		6.662	2.016
Average	322426	28353	161132	101472	966352	0.729	3.178	15.900	23.510
	1130051	88261	621520	813899	3159611	0.172		20.418	4.913

Year	Deposits	Operating	Loans	Securities	Assets	Concent.	ebrd	ir	invgdp
1994	159630	18876	110596	36672	908129	0.724	2.714	36.459	20.384
1995	204555	21161	107384	54715	925280	0.727	2.865	23.979	21.704
1996	219199	23762	103162	57156	832136	0.728	2.865	37.340	22.550
1997	226427	19248	123344	55002	701695	0.731	2.92	18.919	23.363
1998	264755	26558	133510	77402	863296	0.718	3.045	14.577	24.559
1999	291622	23321	142632	129917	918329	0.726	3.197	13.285	23.557
2000	337431	33006	150228	103316	915612	0.739	3.26	12.388	23.870
2001	396119	34002	166735	124261	993573	0.738	3.356	10.864	24.290
2002	416086	30805	192840	142874	975278	0.728	3.423	7.906	23.924
2003	426105	27036	215785	152103	1002068	0.735	3.437	7.163	23.576
2004	487973	39157	222555	170847	1265595	0.733	3.557	6.434	24.736
2005	439210	38217	264817	113394	1340987	0.722	3.666	4.074	25.318

Note: The table presents means and standard deviations of the variables used in the empirical analysis. All bank-level variables are in real terms. Concent. represents a 3-firm concentration ratio derived from own calculations; ebrd the EBRD index of banking sector reform (its standard deviation is not presented); ir a short-term interest rate; and invgdp the ratio of total investment to gross domestic product.

Table 2
Average productive efficiency scores by country and through time

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Average
Bulgaria	0.652	0.701	0.754	0.726	0.686	0.621	0.767	0.727	0.753	0.703	0.681	0.763	0.711
Czech Rep.	0.376	0.490	0.726	0.784	0.663	0.606	0.568	0.451	0.410	0.327	0.340	0.449	0.516
Estonia	0.691	0.716	0.786	0.781	0.707	0.768	0.697	0.789	0.853	0.765	0.830	0.898	0.773
Hungary	0.526	0.541	0.483	0.470	0.623	0.792	0.753	0.692	0.661	0.653	0.658	0.701	0.629
Latvia	0.737	0.836	0.821	0.667	0.578	0.424	0.528	0.570	0.623	0.637	0.707	0.557	0.640
Lithuania	0.751	0.675	0.655	0.382	0.394	0.377	0.565	0.767	0.744	0.853	0.865	0.841	0.656
Poland	0.682	0.663	0.566	0.575	0.680	0.696	0.588	0.487	0.416	0.456	0.489	0.592	0.574
Romania	0.761	0.747	0.766	0.732	0.732	0.757	0.716	0.724	0.796	0.793	0.744	0.825	0.758
Slovakia	0.732	0.773	0.719	0.734	0.751	0.765	0.578	0.534	0.656	0.696	0.718	0.785	0.703
Slovenia	0.580	0.676	0.670	0.605	0.617	0.619	0.731	0.638	0.664	0.596	0.748	0.823	0.664

Table 3
Determinants of bank efficiency using semi-parametric and tobit models

	(1a) SP		(1b) Tobit		(2a) SP		(2b) Tobit		(3a) SP		(3b) Tobit	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficien	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
ebrd	0.2114	7.22	0.1937	6.77	0.1843	6.61	0.1603	5.77				
ref94									-0.1782	-1.22	-0.1054	-0.78
ref95									-0.1663	-1.11	-0.1013	-0.82
ref97									0.2298	4.77	0.2096	5.14
ref98									-0.0748	-1.48	-0.0604	-1.36
ref99									-0.0005	-0.02	0.0041	0.17
ref00									0.1145	2.72	0.0758	2.09
ref01									-0.0393	-1.32	-0.0572	-2.24
ref02									-0.0553	-1.46	-0.0517	-1.62
ref03									0.2663	3.43	0.2533	3.58
ref04									0.1376	2.03	0.0312	1.04
ref05									0.1453	3.49	0.1681	4.63
concentr.	-0.1024	-2.11	-0.0503	-1.06	-0.1122	-1.91	-0.0236	-0.51	-0.1451	-1.98	-0.0251	-0.55
cr					-0.0321	-5.12	-0.0229	-4.49	-0.0367	-5.43	-0.0224	-4.44
size					0.2054	3.70	0.0287	1.14	0.1999	3.64	0.0224	0.98
size squared					-0.0111	-4.96	-0.0065	-1.21	-0.0108	-4.87	-0.0058	-1.07
ir	0.0106	4.42	0.0041	2.88	0.0074	5.03	0.0070	5.10	0.0074	4.89	0.0069	4.71
invgdp	0.1031	2.32	0.0208	0.09	0.0902	1.84	-0.0020	-0.92	0.1093	1.91	0.0031	1.24
for	0.0042	6.11	0.0046	6.43	0.0039	5.98	0.0040	6.79	0.0041	6.20	0.0044	6.39
pub	-0.0022	-3.41	-0.0020	-3.04	-0.0021	-3.31	-0.0023	-3.55	-0.0022	-3.17	-0.0028	-4.08
constant	0.1321	1.34	0.1173	1.04	0.8122	2.21	1.0682	2.99	0.8456	2.31	1.7025	4.62

Note: The table presents estimates of three different specifications using both the Simar and Wilson (SW) and a tobit methodology. The estimations are conducted on an unbalanced data of 4368 observations, corresponding to 364 banks for the years 1994-2005. The dependent variable is bank efficiency as estimated in Stage 1 of our analysis. The explanatory variables are defined as follows: ebrd is the EBRD index of banking sector reform; ref94-ref05 are dummy variables that correspond to a change in the value of ebrd for each country in our sample; concentr. is a 3-firm concentration ratio obtained from own calculations; cr is the ratio of loan loss provisions to loans; size is the natural logarithm of total assets; ir is a short-term interest rate; invgdp is the ratio of total investment to gross domestic product; for is the share of bank assets owned by foreign investors as a % of total assets in the industry; and pub is the share of bank assets owned by the public sector as a % of total assets in the industry.