Intermediation Efficiency of Banks in South-East Europe: An Empirical Assessment Using DEA and Malmquist Index

Valentin Z. Toçi\textsuperscript{a}, Iraj Hashi\textsuperscript{b}

\textsuperscript{a} University of Prishtina, Fakulteti Ekonomik pn, Prishtina, 10000, Republic of Kosovo
\textsuperscript{b} Staffordshire University, Leek Road, Stoke on Trent, ST4 2DF, United Kingdom

Abstract

By employing non-parametric methods, namely Data Envelopment Analysis and Malmquist Index, this paper investigates efficiency of banks using a database of almost all banks in four countries in South-East Europe. The superior efficiency of foreign-owned banks in intermediation is supported. It is argued that the improvement in efficiency of banks originated from the change in technology rather than scale and technical efficiency, and banks on average have not been able to catch-up with best-performers, thus widening the efficiency gap. The largest sources of inefficiency are found to be cost and scale inefficiencies and lending shortfalls. Because of its peculiarities, Kosovo banking sector is assessed relative to other economies. Findings suggest that despite some improvements, the banking system in Kosovo remained less efficient. A number of policy implications emanate from the findings, aiming at enhancing the intermediation efficiency of banks in the context of South-East European transition.

Key words: Bank efficiency; Data Envelopment Analysis; Malmquist Index; South-East Europe

© 2013 Published by SSBFNET

1. Introduction

The relative performance of banks in a market is an essential question for analysis, especially for the South-East Europe (SEE) region that has lagged behind other transition economies, now members of the EU, in banking sector intermediation - despite deep transformation and reforms in the industry undertaken in the first decade. Given that banks are the most important source of external finance for the real sector in SEE, it is vital for policymakers to improve the efficiency of banks in intermediation while maintaining stability. Thus, in this paper intermediation efficiency of banks in four SEE countries is investigated, namely Bulgaria, Croatia, Kosovo and Montenegro in the period 2002-2005.

In the early transition years, SEE countries faced serious challenges in reforming their banking sectors. Subsequently, the states aimed to develop efficient and stable banking systems by allowing foreign banks to enter, on the assumption that this will contribute to efficiency improvements. If foreign banks are more efficient, then an important question to be addressed in this paper is whether there is an overall efficiency improvement in the sector and whether there is any convergence in the efficiency of domestic and foreign banks. Also an important question to be addressed is the intermediation efficiency of banks by the size of the operation, i.e. if inefficiencies are driven by the sub-optimal scale size, in the light of mergers and acquisitions (M&As) of banks in some SEE countries.

This paper puts special emphasis on the case of Kosovo by examining the extent to which the banking sector has been able to catch-up with other countries. Also the period of investigation is suitable for comparisons, given that the banking sector in Kosovo was established from scratch, period 2000-2001 was characterized with new bank entries while in 2005 the banking sector was already an established market. But then underwent structural changes since

\textsuperscript{a} Corresponding author. Valentin Z. Toçi, Tel.: +377 44 376 907.
2006. In 2006 there was one bank bankruptcy and in 2007-2008 there were new entries and some M&As. Therefore, the period 2002-2005 was characterized with a stable structure of the banking sector in Kosovo, making an interesting case for comparing countries with and without stable structures. Furthermore, the global financial crisis that followed staged substantial changes in banking sectors globally, including the SEE region, which is beyond the scope of this study.

The research questions will be explored by employing Data Envelopment Analysis (DEA) and the Malmquist Total Factor Productivity Change Index, which measure the relative efficiency of banks in the SEE from the intermediation point of view using bank-level data. DEA efficiency is estimated for the four SEE countries as a group, several sensitivity tests are employed and the sources of inefficiency are estimated. Also the technical efficiency change, technological change and scale efficiency change on the total factor productivity change is assessed. Finally, some conclusions and policy implications are derived from the analysis.

2. The method

The technical efficiency concept in economics refers to the minimization of inputs used by a firm to produce a given level of outputs or the maximization of outputs produced by a given set of inputs under a given state of technology. However, for any given firm or industry the absolute level of efficiency is unknown. What may be known is the efficiency of a firm relative to another firm or the industry, which is the reference efficiency giving rise to the frontier analysis. The empirical implications of these concepts have been tested by using both econometric and linear programming techniques. Among the non-parametric techniques, formalized first by Charnes, et al. (1978), the DEA is widely used for efficiency measurement in many industries, including banking. A problem with the use of parametric methods in banking is that they require data on input and output prices (cost, revenue or profit function), which are usually not easy to obtain. Apart from data availability, prices derived from a bank’s financial statements are endogenous to the bank’s behaviour leading to the misspecification of the efficient frontier in the presence of market power making the use of price data inadvisable (Pastor et al., 1997; Bos and Kool, 2006). Compared to parametric approaches, an advantage of DEA is that it performs well with small samples, does not require the specification of the functional form and can analyze multiple input-multiple output models, can identify the sources of inefficiency and does not suffer from the possibility of misspecification error and endogeneity. However, the DEA does not take into account any noise and the difference between the efficiency score of an entity and the frontier is all attributed to inefficiency. In this paper, this disadvantage of the method is addressed with some tests such as stability of the efficiency scores over time and the sensitivity of the efficiency scores to outliers.

DEA is a non-parametric linear programming technique used in assessing the relative efficiency of comparable firms such as banks. Banks use some inputs, e.g. labor and capital, in order to produce some outputs, e.g. loans. The DEA method, enables the construction of a frontier as a linear combination of efficient firms with the best combination of inputs and outputs. Any deviation of a particular observation from the efficient frontier is attributed to inefficiency. DEA assumes that all entities under investigation face the same (unspecified) technology, differentials in input costs or product differentiation are not taken into account explicitly, thus, it accounts for technical efficiency. It is important to note that technical inefficiency of financial institutions has generally been found to dominate scale and product mix inefficiencies (Aly et al., 1990; Isik and Hassan, 2002).

The original DEA formulation by Charnes, et al. (1978) assumed constant returns to scale (CRS), i.e. all firms operate at their optimal scale. This assumption has been criticized since, for example, it has been found that in practice the banking sector exhibits non-constant returns to scale (McAllister and McManus, 1993; Wheelock and Wilson, 1999). Factors that may cause banks not to operate at the optimal scale include market power, regulatory requirements, M&As, etc. Banker, et al. (1984) extend the CRS version of DEA to take into account the possibility of variable returns to scale (VRS). The formal treatment of the methodology in a multiple input-multiple output case follows the standard exposition (Cooper et al., 2000; Ramanathan, 2003). Denoting with \( y \) the vector of outputs, \( x \) the vector of inputs, \( v_i \) and \( u_i \) the output and input weights, respectively, \( n \) the number of firms to be evaluated and \( m \) the particular firm whose efficiency is to be measured (\( E_m \)), in the output-oriented case formally we have:

Toçi & Hashi/International Journal of Finance & Banking Studies Vol 2, No 3, 2013  ISSN: 2147-4486
In DEA, multiple inputs and outputs are linearly aggregated using weights and the efficiency of the firm is defined as a ratio of weighted outputs to weighted inputs. DEA assigns a unique (or the best) set of weights to each firm which are determined so as to maximize its efficiency (1), subject to the condition that the efficiency of other firms (calculated using the same set of weights) is constrained to less than or equal to unity (2). This mathematical program produces the efficiency of only one firm, m. For n firms in the sample, n such programs have to be solved. However, this mathematical function has an infinite number of solutions. The problem can be reformulated into a linear programming format by constraining the numerator or the denominator of the efficiency ratio to be equal to unity:

\[
\text{Max} E_m = \frac{\sum_{j=1}^{J} v_{jm} y_{jm}}{\sum_{i=1}^{I} u_{im} x_{im}}
\]

s.t.

\[
\frac{\sum_{j=1}^{J} v_{jm} y_{jm}}{\sum_{i=1}^{I} u_{im} x_{im}} \leq 1
\]

n = 1,2,…,N; v_i, u_i > 0

In the output-oriented DEA, the objective function is to maximize the weighted sum of outputs for firm m (3), while keeping the weighted sum of inputs constrained to unity (4); thus the ratio of weighted outputs to weighted inputs for other firms is constrained to be less than one (5). This is equivalent to the constraint 5, since:

\[
\left[ \sum_{i=m}^{n} \frac{v_{im} y_{im}}{\sum_{i=m}^{n} u_{im} x_{im}} \leq 1 \right] = \left[ \sum_{i=m}^{n} v_{im} y_{im} - \sum_{i=m}^{n} u_{im} x_{im} \leq 0 \right]
\]

This linear programme is an output-maximization DEA programme which implies that other firms cannot produce more output than the reference firm m, if it is efficient. The DEA programmes involving weights for inputs and outputs (v_j and u_i) are called 'multiplier DEA' programmes. Those involving weights for firms (\( \lambda \) – denoting the vector of weights) and the efficiency measure for the firm under evaluation (\( \Theta \)) are 'envelopment DEA' programmes. For each firm under investigation, the linear program finds the
set of weights that maximize efficiency subject to constraints. For the \( m \)-th firm, the general envelopment DEA programme corresponding to the output-oriented CRS model can be presented as follows:

\[
\begin{align*}
\text{Max } & \quad \Theta_m \\
\text{s.t. } & \quad \Theta_m y_{jm} \leq \sum_{j=1}^{J} y_{jn} \lambda_n; \quad j = 1, 2, \ldots, J \\
& \quad x_{im} \geq \sum_{i=1}^{I} x_{in} \lambda_n; \quad i = 1, 2, \ldots, I \\
& \quad \lambda_n \geq 0, \quad n = 1, 2, \ldots, N.
\end{align*}
\]

The objective function is the identification of the maximum feasible expansion of outputs for the firm \( m \), without violating best practice since those firms will have the weighted linear combination of outputs at least as great as that of firm \( m \). So, the firm \( m \) is inefficient if the linear combination of other firms can be found which produce more output (7) and the weighted linear combination of inputs will not exceed the input usage of the firm \( m \) (8). This implies that the output of the firm \( m \) should not exceed the linear combination of other firms, unless it is efficient. If the maximum \( \Theta \) is unity then the firm \( m \) is efficient and will lie on the frontier and its output would equal to that of other firms forming the frontier. If this is the case then \( \Theta_m y_{jm} = \sum_{j=1}^{J} y_{jn} \lambda_n \) and \( x_{im} = \sum_{i=1}^{I} x_{in} \lambda_n \) i.e. no linear combination of firms in a dataset can be found which produce more output levels, by utilizing the same input levels. The firms fulfilling this condition will create the efficiency frontier enveloping all other data points. So \( \lambda \) and \( \Theta \) are variables which need to be calculated, and the linear programming problem must be solved for each firm in order to obtain its respective \( \Theta \). Each efficiency score will be either \( \Theta = 1 \), those units that form the frontier and imply benchmark efficiency, or \( 0 < \Theta < 1 \) implying relative inefficiency compared to the benchmark. Relative efficiency implies that there is no absolute measure of efficiency. Having \( \Theta = 1 \) does not necessarily imply that firms are efficient, but they are not less efficient than other firms in the sample.

As considered above, the CRS model is appropriate when firms operate at the optimal scale. Banker, et al. (1984) proposed an extension of this model to account for variable returns to scale. This necessitates one more restriction for convexity of envelopment surface: \( \sum \lambda_n = 1 \). The introduction of \( \sum \lambda_n = 1 \) in VRS takes into account the variation in efficiency with respect to scale of operation and measures pure technical efficiency. Scale efficiency (Se) can be measured as a ratio of CRS and VRS efficiency scores, with \( Se = 1 \) indicating cases when firms operate at the most efficient scale (size). If \( Se < 1 \) then firms operate at a sub-optimum (‘wrong’) scale. \( 1 - Se \) indicates the relative scale inefficiency (Si) of the particular firm.

DEA measures in input or output orientation estimate the same frontier and the results for efficient entities are the same. For inefficient entities the projected efficiency point on the frontier may differ between input and output orientation. Given that linear programming cannot suffer from statistical problems such as simultaneous equation bias, the choice of an appropriate orientation (input or output) is not a crucial issue as it is in econometric modeling (Fried et al., 1993). Often the choice of orientation depends on the objective of the study. Given the interest in the intermediary role of banks, a more natural choice is the output orientation, i.e. given the deposits, how much increase in lending to the domestic real sector can be attained. In the case of intermediation approach where the volume of deposits is treated as input, conceptually one should choose output orientation. If input orientation is chosen then that
would mean achieving same level of outputs (loans), while minimizing inputs employed (deposits), which is counterintuitive.

Efficiency frontiers are not static over time since the production technology may change, causing positive or negative shifts in the best practice efficiency frontier. DEA results can show that there has been no substantial improvement in the average efficiency in the banking sector, which does not necessarily mean that the productivity has declined. The change in banks’ technology may have led a few banks to take advantage of the new technology which shifted the efficiency frontier outwards while the average bank may have failed to catch-up and take advantage of the new technology, thus, becoming increasingly less efficient relative to frontier banks, though not necessarily less efficient in absolute terms. Hence a 'natural' extension of DEA is the Malmquist Total Factor Productivity Change Index (hereinafter $M_I$). Caves, et al. (1982) introduced the $M_I$ by using distance functions. Extensions to Caves, et al. (1982) entailed the use of DEA in measuring the distance functions and decomposing total factor productivity change into technical efficiency change and change in technology (Färe et al., 1994; Grifell-Tatjé and Lovell, 1997, 1999). Provided that the $M_I$ is obtained by measuring distance functions using DEA, it can be decomposed into four components: technical efficiency change due to VRS assumption, i.e. pure technical efficiency change ($\Delta$PT), scale efficiency change ($\Delta$SE), the product of the two giving technical efficiency change due to the CRS assumption ($\Delta$TE), and the technological change or frontier shift effect ($\Delta$T). The product of all four components gives total factor productivity change ($M_I$). Following Wheelock and Wilson (1999) and Isik and Hassan (2003), the DEA-based Malmquist Total Factor Productivity Change Index ($M_I$) can be decomposed as follows:

$$M_I(X_{t0}, Y_{t0}, X_{t1}, Y_{t1}) = \frac{D_{VRS}(X_{t1}, Y_{t1})}{D_{VRS}(X_{t1}, Y_{t1})} \times \frac{D_{VRS}(X_{t1}, Y_{t1})}{D_{VRS}(X_{t1}, Y_{t1})} \times \frac{D_{CRS}(X_{t1}, Y_{t1})}{D_{CRS}(X_{t1}, Y_{t1})}$$

The first term outside the square brackets ($\Delta$PT) is a measure of pure technical efficiency change between two periods assuming VRS; the first term in square brackets ($\Delta$SE) represents the change in scale efficiency. The product of $\Delta$PT and $\Delta$SE is the overall technical efficiency change ($\Delta$TE). The last term in the square brackets represents the frontier shift effect or change in the technology ($\Delta$T) assuming CRS and the product of all the above gives the $M_I$. For example, if technical efficiency experiences a decline between the two periods or the scale of operation is away from technically optimal scale, then this negatively contributes to the change in $M_I$. $M_I$ and each of its components can be calculated for each adjacent pair of years. For example, as in the case in this paper, for four year period (2002-2005) there will be $M_I$ and its components for three pairs of years.

3. The application of parametric approaches and DEA in measuring bank efficiency in Transition Economies

The literature on the efficiency of banks in transition economies (TEs) is relatively scarce though growing, with the parametric approach dominating the non-parametric approach. In general, parametric studies mainly focus on cost, revenue and/or profit efficiency by assessing the ownership structure of banks and privatization impacts. The studies cover the first decade of transition when many countries faced privatization challenges and the banking crisis were widespread. Policy decisions had to be made in the light of anticipation that the presence of foreign banks would improve the performance of domestic banks (Sabi, 1996). If foreign banks had already entered the market, the question was whether the decision to let foreign banks enter the market was the right one. Table 1 presents the studies, the methodology employed, period of the investigation, number of transition countries included in the analysis and if foreign banks were found to be more efficient. The majority of the studies support what has now become conventional wisdom in transition that foreign banks are more efficient than domestic banks, albeit with notable exceptions. An exception worth discussing is Kraft and Tirtiroglu (1998) who find that state-owned and old private banks in Croatia were more efficient than the newly established private and foreign banks. The authors relate this 'abnormality' to the limited competition and start-up difficulties of the new banks. Although this explanation seems reasonable, it should...
be noted that the period of investigation preceded the crisis of 1998. Systematic misreporting of data by the banks might have been the reason for that outcome. As Grigorian and Manole (2002) have pointed out, the early and mid-transition period was characterized by serious misreporting and non-compliance and this may have been the case in Croatia since the regulatory and supervisory mechanisms, accounting standards and an effective corporate governance framework were not yet in place.

Table 1. Bank efficiency studies in transition employing parametric methods

| Author(s)                     | Methodology | Period of investigation | Number of countries | Foreign banks found to be more efficient than domestic banks |
|-------------------------------|-------------|-------------------------|---------------------|------------------------------------------------------------|
| Kraft and Tirtiroglu (1998)   | SFA         | 1994-95                 | 1                   | no                                                         |
| Kraft, et al. (2002)          | SFA         | 1994-2000               | 1                   | yes                                                        |
| Nikiel and Opiela (2002)      | SFA         | 1993-98                 | 1                   | yes                                                        |
| Hasan and Marton (2003)       | SFA         | 1993-97                 | 1                   | yes                                                        |
| Green, et al. (2003)          | SFA         | 1995-1999               | 9                   | no                                                         |
| Yildirim and Philippatos (2003) | SFA     | 1993-2000               | 12                  | yes                                                        |
| Weill (2003)                  | SFA         | 1997                    | 2                   | yes                                                        |
| Bonin, et al. (2004)          | SFA         | 1996-2000               | 11                  | yes                                                        |
| Matousek (2004)               | DFA         | 1994-2001               | 8                   | yes                                                        |
| Fries and Taci (2004)         | SFA         | 1994-2001               | 15                  | yes                                                        |
| Matousek and Taci (2005)      | DFA         | 1993-1998               | 1                   | yes                                                        |
| Kasman and Yildirim (2006)    | SFA         | 1995-2002               | 8                   | no                                                         |

A potential problem with most of the surveyed studies is their coverage of banks by the dataset. The cross-country studies use the Bankscope database that covers banks with financial statements audited by reputable foreign firms. In some cases banks included in the database represent well below 50 percent of the market. Banks that are mainly larger and sounder tend to be in this database. Consequently, the frontier models in these studies are probably measuring the efficiency of the most efficient banks without indicating the overall position of the sector. In addition, cost or revenue efficiency models that do not account for both (costs and revenues at the same time) may not be appropriate. As Berger and Mester (2003) point out ‘a banker’s decision that raises both costs and revenues, but raises revenues by more than it raises costs, will appropriately be counted as an improvement in performance in profit maximization, but may be counted as deterioration under cost minimization.’ If the intermediation context is added to this observation, e.g. apart from maximizing profits the ability of banks to transform deposits into loans, then a multiple input-multiple output modeling of a banking firm may be more appropriate, which can be handled with non-parametric techniques.

The literature on the application of non-parametric approaches, specifically DEA, in measuring banking efficiency in TEs is scarce. Grigorian and Manole (2002) employ DEA for bank-level data for 17 TEs for the period 1995-1998. The authors found that the privatization of banks, unless controlled by foreign owners, does not result in significant improvements in efficiency. Stavarek (2003) estimates a common efficiency frontier for commercial banks in Poland, Czech Republic, Hungary and Slovakia using DEA for the period 1999-2002, on samples of banks numbering from 59 to 72 in various years. The author found that no improvement in efficiency was evident in the period under review (except in Hungary), greater efficiency is found in the Czech and Hungarian banking sectors and foreign and large banks tended to be more efficient.

Individual country cases applying DEA in measuring bank efficiency are also scarce. Jemric and Vujicic (2002) used DEA to analyze the efficiency of banks in Croatia in the period 1995-2000. Contrary to the findings of Kraft and Tirtiroglu (1998), they found that foreign and new banks were more efficient than domestic and older banks and that a strong equalization in terms of average efficiency of banks had been experienced in Croatia, i.e. there has been a convergence in the sector, which may have been the result of the spillover effects of foreign banks, an increase in competition, etc. The sources of inefficiency in state-owned and old banks were attributed to their large number of employees and large fixed assets. Havrylchyk (2006) investigated bank efficiency using DEA in Poland for 1997-2001 to study the ownership effect and the efficiency improvement of banks over time. His results suggest that foreign
banks were more efficient than domestic banks, though this as a result of foreign banks acquiring the most efficient domestic institutions ('cherry picking') and they have not succeeded in enhancing their efficiency further.

To summarize, efficiency studies in transition yield mixed evidence in relation to the view that foreign banks are more efficient than domestic banks. The foreign bank contribution to the deepening of the financial intermediation in the domestic real sector has not been studied extensively in the multiple input-multiple output modeling using DEA. This is especially the case for the SEE region. This paper contributes to this literature, among others, by analyzing the differences between foreign and domestic banks’ efficiency specifically related to their function as intermediaries, supplemented with the efficiency measurement of the banking industry by size of operation, cross-country differences, etc.

4. The definition of inputs and outputs

Selecting appropriate input and output variables is perhaps the most important step in using DEA since it determines the evaluating context of the comparison (Yeh, 1996). There is a divergence and remains an open question in the literature about the correct definition of inputs and outputs in banking. There is some agreement that loans and other interest bearing assets should be treated as outputs. However, there is a disagreement on the classification of deposits as the main ingredient for loan production and other earning assets (an input) or as a service provided to clients and resource consuming activity (an output).

In general, there are some approaches that most studies use and, to some extent, the choice may be settled depending on which approach is adopted. For example, the production approach views banks as producers of services for depositors and borrowers, e.g. administering customer accounts and transactions, cashing cheques and issuing loans, while using some combination of labor and capital as inputs. A variant of the production approach, known in the literature as the value-added approach, treats as bank outputs those activities that have value added in the sense that they incur significant labor and capital costs. As a result, both asset and liability items may have input and output characteristics and outputs are defined on the basis of the resource consuming activity that a particular bank product represents.

The intermediation approach, pioneered by Sealey and Lindley (1977), views banks as intermediaries between the savers and borrowers, considering deposits as inputs which are the source of bank lending. It ignores the output characteristics of deposits in providing liquidity and payment services to customers, and as a resource consuming activity in the banking business. A variant of the intermediation approach is the asset approach which treats loans and other interest bearing assets of banks as their main outputs. Another variant, the profit approach, is concerned with profit maximization where all cost and revenue components are decision variables to be optimized by bank managers (Milima and Hjalmarsson, 2002).

The risk-management approach is a departure from the intermediation approach, incorporating risk management, information processing and monitoring activities of the bank (Colwell and Davis, 1992; Mester, 1996). Risk management characteristics were often ignored in previous studies. For example, banks could boost their lending activity by extending loans on unsound base resulting in bank instability and failure. Thus, one of the missing ingredients in the intermediation approach is covered by the risk-management approach by introducing the quality of bank assets. If available, the data on NPLs can be used to correct for overall lending activity of the bank. However, usually the data on NPLs are difficult to obtain and approximations such as loans net of provisioning are considered.

The ability of banks to transform deposits into loans directed towards domestic real sector emphasizes the intermediary role of banks and as such is the main motivation for following the intermediation approach in this paper. The variables consist of two inputs and two outputs with deposits and total costs being considered as the bank inputs, and loans net of provisions and total revenues as outputs. Total costs include interest expenditures and non-interest expenditures such as salaries, expenditures related to fixed assets, administrative expenses and similar, excluding taxes to avoid any distortion. Total revenues include interest income and non-interest income such as fees and commission income, income from the trading portfolio, etc. Unlike some studies which take interest and non-interest revenues and expenditures separately, the data are aggregated into the total cost and total revenues for several reasons.

First, it is debatable to what extent the data on inputs (or outputs) should be disaggregated. Apart from the unsettled problem of the input/output definition in banking, there is also the problem of the disaggregating measures used. On the input side, just to mention a few, there are labor and capital inputs. Banks spend resources on personnel which can be disaggregated into clerical and managerial. The capital input is usually taken to be a fixed asset (the balance sheet
or income statement item), although banks also spend resources on information technology and marketing which are important parts of the cost structure in modern banking. On the output side, e.g. loans, the list of disaggregation may also be long. For example, loans may be broken down by economic sector, maturity, type of borrowers and all these may be considered as different products given their different input usage, specific risks, etc.

The degree of disaggregation of data also depends on the aims of the researcher. If the precise source of inefficiency is to be captured, then as detailed a breakdown as possible would be necessary. For example, disaggregating between labor and fixed assets would be appropriate for the management of the bank, e.g. if the source of the inefficiency is overstaffing, then this information may be very important for improving efficiency in this respect. However, if the interest is in more general characteristics of the banking sector, as it is the case in this paper, i.e. how the ownership structure and size (foreign vs. domestic banks; large vs. small banks) affect efficiency, then the detailed breakdown, even though informative, would not be of much of relevance. As a result, on the input side, total costs would capture all the operational spectrum of banking activity including staff, marketing, information technology, interest expenditures, etc.

To some extent this choice of inputs and outputs in this paper may be considered as a 'mixed approach' since some important elements of all approaches are taken into account.

First, it is a version of the intermediation approach since deposits are defined as inputs and the role of banks in the intermediation process is captured by loans treated as outputs.

Second, to control for banks undertaking too much risk – the risk-management approach – net loans (after provisions) are used, which assumes appropriate provisioning by banks which in turn depends on the supervisory capacity of the regulator. Given that excessive risk taking would imply a waste of resources and disintermediation rather than intermediation, the use of loans net of provisioning will take into account risk preferences. If under-provisioning is an issue, then this would not be captured by the data given that data on the quality of assets are not available. However, to some extent this is also picked up in the interest income (as part of total revenue) which would capture the actual returns on loans.

Third, by having total costs as an input and total revenues as an output also some elements of the profit approach are taken on board. Berger and Mester (2003) argue that the use of the profit approach may help take into account unmeasured changes in the quality of banking services by including the higher revenues paid for improved quality.

Fourth, given the focus on the efficiency of banking sector in intermediating funds to the real sector, only loans to the domestic real sector are considered as an output and assets held in government securities, banks abroad, and inter-bank loans are excluded. Where banks invest a significant part of their assets in low risk low-return foreign or domestic government securities, from the intermediation point of view it is preferable not to treat this part of assets as bank output. Also from the intermediation point of view non-loan assets incur little screening and monitoring costs and are excluded from the measures of financial development at the macroeconomic level. So in this context, net loans capture the output of banks from the intermediation point of view.

Finally, the output nature of deposits (and non-loan earning assets) would be captured by the income from fees and commissions and interest income, respectively, which are incorporated in the total revenue variable. Hence, having total revenues on the output side would account for the output nature of banking services. Interest and non-interest expenses/revenues would capture the entire spectrum of banking activities, including off-balance sheet items, the lack of account of which is a criticism of the intermediation approach. Therefore, the ability of banks to transform deposits into loans and the ability to maximize revenues at given costs (and consequently better profits) would imply efficient intermediation.

5. The intermediation efficiency of banks in SEE: empirical results

In this section, both the CRS and VRS models are used to analyze the relative efficiency of banks in SEE. The common efficiency frontier for banks in the four SEE countries is estimated. Relative efficiency of banks for individual countries is assessed as well and results are broadly similar compared to the common frontier approach. Some important findings for individual countries are briefly discussed (the results for individual country DEA scores are available upon request to the authors). The output oriented multi-stage DEA model is utilized, using the DEAP 2.1 software developed by Coelli (1996). The data for each bank are taken from the end-of-year balance sheet and income statements published by central banks in the respective countries.
The database virtually covers the entire banking system in Bulgaria, Croatia, Kosovo and Montenegro. Strength of having the entire population of banks is that it would not suffer from the problem of DEA being a sample specific method, since the results obtained for a sample of banks may not be generalized for the entire population. For Bulgaria and Croatia the data are in local currencies and the official exchange rate is used to convert them to euro, and for Kosovo and Montenegro they are in euro. The dataset contains 86 banks in 2002, 83 in 2003, 78 in 2004 and 84 in 2005. The sample size is increased in 2005 due to the inclusion of 10 banks from Montenegro. While the number of banks remained the same in Kosovo throughout the period, they continuously declined in Bulgaria and Croatia due to M&As and some bank closures.

At the outset it should be noted that banking sectors in SEE countries continued the consolidation process in the 2002-2005 period. This is especially the case with the banking sector in Croatia where the consolidation process involved 19 banks in 12 M&As, and three bankruptcies. It is important to note that almost all banks involved in the consolidation process were foreign-owned. In Bulgaria the consolidation was less pronounced. It involved five banks undergoing M&As, three new entries and two exits. Usually motives for consolidation are attributed to the opportunities in exploiting economies of scale, such as spreading costs over a larger base, economies of scope, such as increasing the number of services offered, cross-selling, increasing customer base, etc. Other motives involve risk diversification, acquisition of the new technology or an increase in the market power – the latter may negatively affect efficiency.

Two questions related to M&As will be addressed explicitly in this paper: (i) if scale efficiency improved and, (ii) if a shift in technology appeared in this period. On the other hand, Kosovo had stable structure of the banking sector in this period making an interesting case for comparison; while the data for Montenegro are available only for one year. The development of the banking sector in Kosovo is one of substantial expansion in all dimensions, although from a low base. From the regulatory point of view, it is worth mentioning that in Kosovo the upper limit of loan to deposit ratio is at 80 percent, in Croatia the reserve requirements increased several times in the 2002-2005 period in an attempt to curb the rapid credit expansion while Bulgaria introduced such measures in 2005. These measures also may have had an impact on the intermediation efficiency of banks in the region.

5.1. A common frontier of banks in SEE

In this section the efficiency of the SEE banking systems over the period 2002-2005 is assessed. Constructing a common frontier for different countries may pose some problems because of differences in macroeconomic, regulatory environment and accounting standards. However, there are justifications for considering that countries may be treated as largely homogenous because the period under review is characterized by a stable macroeconomic environment in the countries under investigation. Kosovo and Montenegro have no independent monetary policy as they are euroised economies while Bulgaria is quasi euroised with a currency board. Croatia is also highly euroised and loans and deposits are largely indexed to the euro. Foreign banks dominate all banking markets in the region. From the regulatory point of view all the countries have 'deregulated' markets, i.e. there are no controls on interest rates, capital flows, etc. Regarding the differences in accounting standards, the problem is attempted to be minimized by adopting broad definitions of inputs and outputs. Given the above, the procedure may be warranted.

Table 2 presents the evolution of banking sector intermediation efficiency in the four countries. The average efficiency scores in the CRS model for the whole region fluctuate in the four years reaching their minimum in 2003 (0.679) and maximum in 2004 (0.714). In general, between 2002 and 2005, there seems to have been some marginal improvement in the efficiency of banks in the region, although there have been no efficiency gains.

The divergence among banks, as measured by the standard deviation, seems not to be negligible and varying yearly. In estimates for individual country cases, DEA scores show that the divergence between best and worst performing banks in the market is widening; foreign/larger banks are more efficient than domestic/small banks and the main source of inefficiency emanates from scale inefficiencies where most of the banks operate at decreasing returns to scale (DRS).

As can be seen, Bulgarian banks are, on average, more efficient in intermediation than Croatian banks, while banks in Kosovo appears to be the least efficient compared to other countries in the region. DEA scores for Bulgaria only show that there was a drop in efficiency in Bulgarian banking sector in the period under review, mainly attributed to domestic banks and introduction to supplementary reserve requirements by the central bank to slow down rapid credit growth. Similar results seem to hold in 2005 when banks from Montenegro are included. Nonetheless, Montenegrin banks appear to be the most efficient ones, on average. The difference between the average of the least efficient banking system (Kosovo) and the most efficient system (Montenegro) in 2005 is around 22.1 percentage points. The
efficiency spreads between Bulgaria and Croatia are not as large (varying between three to five percentage points in different years). In general, the VRS model tells the same story. An exception is that in 2002 banks in Croatia are more efficient than banks in Bulgaria, pointing to the scale inefficiencies in the Croatian banking sector.

Regarding the breakdown of banks by ownership, the results suggest that foreign banks on average outperform domestic banks in terms of average efficiency and the number of banks dominating the frontier and the gap was increasing over the years.

Table 2. DEA efficiency scores of banks in SEE (2002-2005), CRS and VRS model

| Country                     | CRS       | VRS       |
|-----------------------------|-----------|-----------|
|                             | 2002  | 2003     | 2004     | 2005 | 2002 | 2003 | 2004 | 2005 |
| Number of banks (of which foreign-owned) | 86(44) | 83(41) | 78(39) | 84(38) | 86(44) | 83(41) | 78(39) | 84(38) |
| Number of efficient banks (of which foreign-owned) | 5 (3) | 7(5) | 7(5) | 6(4) | 15(10) | 18(13) | 17(14) | 17(14) |
| Average efficiency          | 0.684 | 0.679 | 0.714 | 0.693 | 0.812 | 0.780 | 0.791 | 0.778 |
| Standard deviation          | 0.153 | 0.169 | 0.144 | 0.151 | 0.150 | 0.161 | 0.156 | 0.163 |
| Average efficiency of banks in Bulgarla | 0.699 | 0.708 | 0.751 | 0.708 | 0.818 | 0.808 | 0.822 | 0.789 |
| Average efficiency of banks in Croatia | 0.696 | 0.676 | 0.698 | 0.664 | 0.843 | 0.787 | 0.789 | 0.772 |
| Average efficiency of banks in Kosova | 0.532 | 0.556 | 0.618 | 0.593 | 0.574 | 0.604 | 0.654 | 0.659 |
| Average efficiency of banks in Montenegro | na | na | na | 0.814 | na | na | na | 0.844 |
| Average efficiency of foreign banks | 0.693 | 0.695 | 0.736 | 0.725 | 0.827 | 0.825 | 0.842 | 0.824 |
| Average efficiency of domestic banks | 0.672 | 0.658 | 0.687 | 0.649 | 0.792 | 0.720 | 0.727 | 0.714 |
| Efficiency Gap (Foreign - Domestic) | 0.021 | 0.037 | 0.049 | 0.076 | 0.035 | 0.105 | 0.115 | 0.110 |
| Ownership                   |          |          |
| Average efficiency of foreign banks |          |          |
| Average efficiency of domestic banks |          |          |
| Efficiency Gap (Foreign - Domestic) |          |          |
| Size                        |          |          |
| First Quartile              | 0.700 | 0.718 | 0.748 | 0.718 | 0.796 | 0.755 | 0.778 | 0.764 |
| Second Quartile             | 0.633 | 0.608 | 0.652 | 0.618 | 0.766 | 0.727 | 0.711 | 0.698 |
| Third Quartile              | 0.673 | 0.705 | 0.774 | 0.760 | 0.876 | 0.917 | 0.929 | 0.930 |
| Fourth Quartile             | 0.795 | 0.631 | 0.661 | 0.707 | 0.968 | 0.938 | 0.961 | 0.959 |

To explore further the differences in efficiency scores between foreign and domestic banks statistically the non-parametric Mann-Whitney Rank-Sum test is used, as suggested by Cooper, et al. (2000), to test the null hypothesis that foreign and domestic banks are drawn from a population having the same distribution, i.e. whether the differences between the two groups are statistically significant. The results are presented in Table 3 for individual country cases and the SEE region as a whole.

As can be observed from the table, the differences between efficiency scores of foreign and domestic banks are statistically significant at the five percent level in Bulgaria in 2003 and after (and at the ten percent level in the VRS model in 2005). In Croatia the differences are statistically significant only in 2005 at the ten percent level of significance, while in Kosovo the differences are statistically significant in favor of domestic banks. This is reversed in 2004 in the VRS model only. In 2005, the differences between the efficiency scores of foreign and domestic banks are not significant in Kosovo or Montenegro. In the common sample of efficiency scores for banks in SEE countries the results indicate that the differences between foreign banks and domestic banks are more apparent only after 2003 and
it seems that the results are mainly driven by the difference between foreign banks and domestic banks in Bulgaria and in 2005 for Croatia, e.g. after the consolidation took place.

|                | 2002   | 2003     | 2004     | 2005     |
|----------------|--------|----------|----------|----------|
| **Bulgaria**   |        |          |          |          |
| CRS (z-stat)   | 1.135  | 2.399**  | 2.521**  | 2.145**  |
| VRS (z-stat)   | 0.510  | 2.241**  | 2.148**  | 1.590*   |
| **Croatia**    |        |          |          |          |
| CRS (z-stat)   | 0.594  | 0.209    | 1.478    | 1.614*   |
| VRS (z-stat)   | 0.608  | 0.478    | 1.294    | 1.878*   |
| **Kosova**     |        |          |          |          |
| CRS (z-stat)   | -6.898*** | -6.061*** | -1.784*  | 0.891    |
| VRS (z-stat)   | -2.221*** | 0.164    | 2.612***  | 0.478    |
| **Montenegro** |        |          |          |          |
| CRS (z-stat)   | na     | na       | na       | -0.365   |
| VRS (z-stat)   | na     | na       | na       | 0.125    |
| **SEE**        |        |          |          |          |
| CRS (z-stat)   | 0.683  | 1.116    | 1.591*   | 2.430***  |
| VRS (z-stat)   | 1.260  | 2.856*** | 3.218***  | 2.973***  |

Note: ***, **, * denote significance at 1%, 5% and 10%, respectively.

Regarding the difference in efficiency scores between banks of different size classes the evidence is not so clear (Table 2). For example, in the CRS model small banks in the first quartile are more efficient than larger banks (except in 2002 compared to fourth quartile and 2004-2005 compared to third quartile); while in the VRS model banks in third and fourth quartile appear the most efficient. If we recall that CRS is a product of technical and scale efficiency while VRS is a measure of pure technical efficiency, with any difference between the two reflecting scale inefficiencies. Table 4 presents scale inefficiencies of banks in SEE. As can be seen, while scale inefficiencies are not substantial for small banks, they appear to be considerable for large banks which generally operate under DRS – a greater efficiency would have been attained had they operated at lower scale. Banks in Croatia have the largest scale inefficiencies, followed by banks in Bulgaria. Also foreign banks seem to have higher scale inefficiencies than domestic banks. The consolidation of the banking sectors in Croatia, and to some extent Bulgaria, may be an explanation for this outcome.

In general, the evidence suggests that economies of scale are exploited early in banking in terms of the size and that there is a general consensus that consolidation through M&As is beneficial up to a certain size, relatively small, in order to reap economies of scale (see for example, Ferrier and Lovell, 1990, McAllister and McManus, 1993, Amel et al., 2004). Consequently, the size of banks (in terms of total assets) exhibiting constant returns to scale (CRS) in Bulgaria and Croatia is explored. The average size of banks exhibiting CRS for the 2002-2005 period in Bulgaria was €116 million (minimum €63 million in 2002 and maximum €130 million in 2005). In Croatia the average size of banks exhibiting CRS was 141 million (minimum €20 million in 2002 and maximum €300 million in 2005), thus, confirming to some extent previous assertions that the scale economies in banking are not so large. However, there are two banks in Bulgaria with the asset size of around €400 million and two in Croatia with the asset size of around €1 billion (one of which was not involved in M&A) that exhibit CRS. These exceptional cases make it difficult to specify the optimal scale size of banks since it depends on the characteristics of the particular market.

While the optimal scale of operation in banking is difficult to ascertain, the results suggest that the greatest source of inefficiency in the SEE banks is scale inefficiency of large banks attributable to M&As. It should be noted, however, that it takes time for the positive effects of M&A to take place. This is because restructuring and integrating the acquired bank involves changes in organizational culture, operations, policies and procedures, and staff training to adapt to the new setting. It also involves branch and staff rationalization, staff turnover, change in cultures which may cause losses in customer relationships. From the results in the period under review it may be said that little gain in efficiency would be attained through M&A, especially for large banks, given they are generally operating at DRS.
The period under review in SEE is characterized by a growing economy and an increase in market demand that may have encouraged the banks’ excessive size. It appears that some banks attempted to become larger for positioning in the market and increase the market share, to preserve margins at the expense of scale economies. However, another motive may have been the acquisition of new technologies (explored in Section 7) which may be reflected in cost inefficiencies (Section 6.3).

Several observations can be made so far. In the banking sectors in the SEE countries no substantial efficiency improvements are apparent and there seems to have no convergence in the banking market in terms of intermediation efficiency. A notable exception is Kosovo which is showing that the catching-up with other countries is maybe taking place, although still remaining the least efficient banking sector in the region. In general, foreign banks appear more efficient than domestic banks and the differences are widening. Regarding the performance of banks of different size, large banks are more efficient than small banks, though large banks suffer from scale inefficiencies and banking markets are characterized with excessive size.

Table 4. Scale inefficiencies of banks in SEE, 2002-2005

| Country       | 2002   | 2003   | 2004   | 2005   |
|---------------|--------|--------|--------|--------|
| Average       | 0.157  | 0.130  | 0.097  | 0.109  |
| Bulgaria      | 0.145  | 0.124  | 0.086  | 0.103  |
| Croatia       | 0.174  | 0.141  | 0.115  | 0.140  |
| Kosova        | 0.073  | 0.079  | 0.055  | 0.100  |
| Montenegro    | na     | na     | na     | 0.036  |

Ownership

| Ownership     | 2002   | 2003   | 2004   | 2005   |
|---------------|--------|--------|--------|--------|
| Foreign banks | 0.162  | 0.158  | 0.126  | 0.120  |
| Domestic banks| 0.152  | 0.086  | 0.055  | 0.091  |
| Scale inefficiency Gap (Foreign - Domestic) | 0.011 | 0.071 | 0.071 | 0.029 |

Size

| Size          | 2002   | 2003   | 2004   | 2005   |
|---------------|--------|--------|--------|--------|
| First Quartile| 0.121  | 0.049  | 0.039  | 0.060  |
| Second Quartile| 0.174  | 0.164  | 0.083  | 0.115  |
| Third Quartile| 0.232  | 0.231  | 0.167  | 0.183  |
| Fourth Quartile| 0.179  | 0.327  | 0.312  | 0.263  |

Note: Scale inefficiency = 1 - [(CRS/VRS)]

6. Sensitivity tests

In this section, the sensitivity of the DEA results is tested in three ways. First, whether the constructed frontier is influenced by outliers; second, whether the efficiency scores are stable over time; and third, whether the DEA scores are consistent with financial ratio analysis.

6.1. Test for outliers

As shown in the previous section, in the CRS model the efficiency frontier was determined by five to seven banks throughout the period. One possible problem with DEA is that the frontier may be defined by the outliers rather than the whole sample, e.g. because of data errors (Simar, 2003) and as such be susceptible to extreme observations. The results of the test for outliers are displayed in Table 5.
Table 5. Spearman rank-order correlation coefficients: test for outliers

| Year | No. of efficient banks (original sample) | No. of efficient banks (reduced sample) | Rank-order correlation (all sample vs. reduced sample) |
|------|----------------------------------------|----------------------------------------|------------------------------------------------------|
| 2002 | 5                                      | 7                                      | 0.8867***                                              |
| 2003 | 7                                      | 11                                     | 0.9397***                                              |
| 2004 | 7                                      | 10                                     | 0.8645***                                              |
| 2005 | 6                                      | 10                                     | 0.9830***                                              |

Note: ***, **, * denote significance at 1%, 5% and 10% level, respectively.

The first column shows the number of efficient banks that construct the frontier in the original sample. Next, these banks are deleted from the sample and the efficiency scores are re-estimated for the reduced sample (the number of efficient banks for the reduced sample is shown in the second column). Then the Spearman rank-order correlation is calculated between the two estimates of efficiency scores in the two samples and it is found that correlation coefficients are high (ranging from 0.86 to 0.98) and are statistically significant at the one percent level. Similarly, high correlation coefficients have been found in Berg, et al. (1993) ranging from 0.95 to 0.99 employing a similar procedure. This may indicate that the results are not sensitive to outliers, i.e. those banks that appeared as more efficient in the whole sample were likely to remain such in the reduced sample.

6.2. Test for stability over time

A test for the stability of efficiency scores over time is conducted since, as pointed out by Leong, et al. (2003), banks cannot genuinely be efficient in one year and then highly inefficient in the next year, except in extraordinary circumstances. If the DEA scores are stable over time, then the problem of noise due to measurement errors, luck and lack of comparable entities for banks dominating the frontier diminishes. Therefore, at least within a short period of time, it is desirable that the efficiency scores are reasonably stable. Leong, et al. (2003) find low stability of DEA efficiency scores over time in their study of the banking sector in Singapore (correlation coefficients ranging from 0.14 to 0.06 within a three year time span). On the other hand, Bauer, et al. (1998) and Fiorentino, et al. (2006) in their studies of the banking sector in the U.S. and Germany, respectively, find quite large correlation coefficients ranging from 0.92 to 0.76 in a three year time span, suggesting that relative repositioning in the banking industry neither appears to occur quickly nor to a large extent.

Table 6. Spearman rank-order correlation coefficients of efficiency measures over time

|                  | One year apart (2004) | Two years apart (2003) | Three years apart (2002) | Five years apart (2005) |
|------------------|-----------------------|------------------------|--------------------------|-------------------------|
| CRS 2005         | 0.7875***             | 0.6905***              | 0.6276***                | 0.6882***               |
| VRS 2005         | 0.8717***             | 0.7887***              | 0.6834***                | 0.7588***               |

Note: ***, **, * denote significance at 1%, 5% and 10%, respectively.

Table 6 contains the Spearman rank-order correlation coefficients which represent the correlation of efficiency scores in 2005 with the scores in 2004, 2003 and 2002. As the results indicate, the coefficients are quite high and significant at the one percent level suggesting that many of the worst performers and best performers have tended to remain such over time. For example, the rank correlation coefficients for the CRS and VRS efficiency scores in 2005 with the CRS and VRS scores one year forward are 0.79 and 0.87, respectively. The correlation coefficients decrease over time as one would expect, but remain significant at one percent level suggesting that the model ranks banks fairly consistently over time.
6.3. DEA, Financial Ratios and Sources of Inefficiency

Financial ratio analysis is often used by the bank management, regulators, financial analysts as well as researchers to evaluate bank performance based on some 'benchmarks' for the industry. One disadvantage of the ratio analysis is that each ratio must be compared with some benchmark ratios, one at a time, while assuming that other factors are fixed and that the benchmarks chosen are suitable for the purpose of comparison (Yeh, 1996). Among the most frequently used accounting ratios are return on assets (ROA) and return on equity (ROE) which proxy for bank efficiency in generating profits; cost to total asset ratio (CTA), cost to income ratio (C/I) and revenue to asset ratio (RTA) which are indicators of optimization in terms of bank costs and revenues; loans to total assets (LTA) and loans to total deposits (L/D) which show the extent to which assets/deposits are devoted to loans as opposed to other assets; capital to total assets ratio (CAR) which shows capital adequacy, etc. A desirable property of the DEA model is its usefulness for policy purposes in serving regulators in identifying the worst performing banks in the market. In banking studies, apart from the extensive use in efficiency measurement, DEA has also been proposed for use in monitoring and/or early warning systems by bank regulatory agencies since it has a strong empirical association with bank failures after the inefficiencies have been detected (Barr et al., 1994; Kao and Liu, 2004). As Bauer, et al. (1998) pointed out, ‘a positive correlation between efficiency scores (generated in DEA or parametric methods) and the usual ratios would be desirable because the authorities could be more confident that the measured efficiencies are accurate indicators of performance and not just artefact of the assumptions made regarding the underlying optimization concept, the inexistence of random error and, for parametric methods, the functional form assumptions and distributional assumptions of efficiency scores and the random error.’

Table 7. Spearman rank-order correlation coefficients of efficiency measures and financial ratios

|       | ROA      | ROE      | NIM      | C/I     | CTA     | RTA     | LTA     | L/D     | CAR     |
|-------|----------|----------|----------|---------|---------|---------|---------|---------|---------|
| CRS   |          |          |          |         |         |         |         |         |         |
| 2002  | 0.3882***| 0.2936** | 0.4873***| -0.4409***| -0.0126 | 0.2592**| 0.6846***| 0.8311***| 0.2912**|
| 2003  | 0.3195**  | 0.0211   | 0.4445***| -0.4705***| -0.0688 | 0.4434***| 0.7439***| 0.8735***| 0.5007***|
| 2004  | 0.4106***| 0.0842   | 0.4853***| -0.5323***| -0.0004 | 0.3907***| 0.6064***| 0.7555***| 0.4710***|
| 2005  | 0.4845***| 0.2074   | 0.4618***| -0.6110***| -0.1622 | 0.3600***| 0.5851***| 0.7873***| 0.4555** |
| VRS   |          |          |          |         |         |         |         |         |         |
| 2002  | 0.4203***| 0.3294***| 0.2925** | -0.4505***| -0.0694 | 0.2089* | 0.3932***| 0.5159***| 0.2170*  |
| 2003  | 0.3515***| 0.2759** | 0.2139*  | -0.5300***| -0.2285*| 0.1782  | 0.4929***| 0.5528***| 0.1445   |
| 2004  | 0.3381***| 0.2526** | 0.1857   | -0.5528***| -0.2967**| 0.0441  | 0.5075***| 0.5541***| 0.1803   |
| 2005  | 0.4483***| 0.3410***| 0.3024** | -0.6092***| -0.3023**| 0.1751  | 0.5904***| 0.7155***| 0.2693** |

Note: ***, **, * denote significance at 1%, 5% and 10%, respectively; ROA (return on assets) is the ratio of pre-tax profit to total assets; ROE (return on equity) is the ratio of pre-tax profit to total equity; NIM (net interest margin) is the difference between the interest income and interest expenditures divided by total assets; C/I is the ratio of total costs to total income; CTA is the ratio of total costs to total assets; RTA is the ratio of total revenues to total assets; CAR is the ratio of loans to total assets; L/D is the ratio of loans to deposits; ROA (return on assets) is the ratio of pre-tax profit to total assets; ROE (return on equity) is the ratio of pre-tax profit to total equity; NIM (net interest margin) is the difference between the interest income and interest expenditures divided by total assets; C/I is the ratio of total costs to total income; CTA is the ratio of total costs to total assets; RTA is the ratio of total revenues to total assets; CAR is the ratio of loans to total assets; L/D is the ratio of loans to deposits; CAR (capital adequacy ratio) is the ratio of total equity to total assets.

Table 7 presents the correlation coefficients of DEA efficiency scores with some broadly used financial ratios. As indicated in the table, the correlation coefficients are, in most cases, reasonably large and statistically significant. The largest correlation coefficients are exhibited with loan to deposit ratio (L/D) ranging from 0.52 to 0.87, loan to total assets ratio (LTA) ranging from 0.39 to 0.74, cost efficiency ratio (C/I) ranging from 0.44 to 0.61 – all statistically significant at one percent level – followed by profit to total assets ratio (ROA) ranging from 0.32 to 0.48. These correlation coefficients are higher than the ones reported in other studies. For example, Isik and Hassan (2002) find
correlation coefficients of efficiency scores with ROA, CTA and C/I ratios in range 0.30-0.40, Bauer, et al. (1998) in range 0.10-0.20, while Fiorentino, et. al. (2006) find even weaker correlation coefficients.

Table 8. Statistical tests of differences in financial ratios for the best and the worst DEA performers (2005)

| CRS          | ROA   | ROE   | NIM  | C/I  | CTA  | RTA  | LTA  | L/D  | CAR  |
|--------------|-------|-------|------|------|------|------|------|------|------|
| Banks (DEA=1) | 3.78  | 20.77 | 5.82 | 60.76 | 5.70 | 9.79 | 68.96 | 91.60 | 29.77 |
| Banks (DEA<Average DEA) | 0.89  | 1.52  | 3.57 | 84.93 | 6.98 | 8.23 | 47.06 | 54.91 | 10.84 |

Mann-Whitney Test

| z-stat | 2.331** | 2.310** | 1.626 | -2.923*** | -1.367 | 0.849 | 2.310** | 3.441*** | 3.365*** |
|--------|---------|---------|-------|------------|--------|-------|---------|----------|----------|

Table 8. Statistical tests of differences in financial ratios for the best and the worst DEA performers (2005)

| VRS          | ROA   | ROE   | NIM  | C/I  | CTA  | RTA  | LTA  | L/D  | CAR  |
|--------------|-------|-------|------|------|------|------|------|------|------|
| Banks (DEA=1) | 2.51  | 17.92 | 4.48 | 66.3 | 5.7  | 8.71 | 61.22 | 77.77 | 17.59 |
| Banks (DEA<Average DEA) | 0.89  | 0.92  | 3.74 | 85.82 | 7.74 | 8.69 | 48.28 | 56.85 | 11.42 |

Mann-Whitney Test

| z-stat | 2.409** | 1.880* | 0.740 | -3.791*** | -3.041*** | 0.454 | 3.060*** | 3.791*** | 1.630 |
|--------|---------|--------|-------|------------|------------|-------|----------|----------|-------|

Note: ***, ***, * denote significance at 1%, 5% and 10%, respectively; ROA (return on assets) is the ratio of pre-tax profit to total assets; ROE (return on equity) is the ratio of pre-tax profit to total equity; NIM (net interest margin) is the difference between the interest income and interest expenditures divided by total assets; C/I is the ratio of total costs to total income; CTA is the ratio of total costs to total assets; RTA is the ratio of total revenues to total assets; LTA is the ratio of loans to total assets; L/D is the ratio of loans to deposits; CAR (capital adequacy ratio) is the ratio of total equity to total assets.

Further, the average of selected ratios for the most efficient banks identified by DEA in the CRS and the VRS model are computed, i.e. frontier banks, and below average performers in the market, i.e. banks with below average DEA efficiency scores. As can be seen from the Table 8, on average, better DEA performers have also better financial ratios and the differences seem quite large. For example, more efficient banks have larger return indicators, are better in managing costs, are better capitalized, and are better in transforming assets/deposits into loans. Given the small sample size, the non-parametric Mann-Whitney test is conducted in order to draw inference on the differences in accounting ratios between the 'best' and the 'worst' DEA performers. As can be observed from Table 8, in the both CRS and VRS models, significant differences appear for almost all the accounting ratios.

The above results may suggest that inefficiencies originate in the process of transforming deposits into loans and extending loans as opposed to other assets as well as cost inefficiencies. However, as discussed earlier in this paper, the DEA method has the advantage of identifying the source of inefficiency directly. By constructing the frontier of best performers, best achievable output and input targets are computed. The difference between the actual outputs/inputs of inefficient units and outputs/inputs of best practice units may represent the sources of inefficiency. Calculating the mean values of output shortfalls for the sector and mean values of the excess use of inputs, and relating this to the mean value of actual outputs and inputs gives the value of inefficiencies for each input and output. By calculating this, results suggest that the greatest source of inefficiency are excess costs (2.60 percent), followed by a shortfall in loan extension (1.31 percent) – broadly in line with the ratio analysis. In all, the results show that the DEA model efficiency scores are comparable with most of the standard performance measures, indicating that using DEA in conjunction with the usual ratio analysis may be a useful tool.

7. Total factor productivity change

Table 9 presents the results of the Malmquist Index (M) for the banking sector in SEE in the period 2002-2005. As can be seen, it appears that there was a minor improvement in M (0.8 percent) in the SEE countries in terms of banking sector intermediation, driven by the change in technology (∆T) rather than the improvement in technical (∆PTE and ∆TE) and scale efficiencies (∆SE) – which generally exhibit efficiency decline. The technology shift effect
improved by 3.2 percent (ΔT), which is offset by the catching-up effect with the overall decline of 2.3 percent (ΔTE). The change in technology is positive in two years while negative in one and the opposite is true for other components.

Technological progress characterized banks of all classes (by ownership, size and country), although banks in Kosovo, smaller and domestic banks showed lower progress. In terms of M_t, however, the largest progress is apparent for the banks in Kosovo in all of its components in the period under review. This shows that Kosovo was catching-up with other countries in terms of financial intermediation. Croatia exhibits a declining M_t in the period under review and the primary source is the decline in the technical and scale efficiency, while there is a positive shift in technology. Bulgaria, on the other hand, exhibits a small improvement in M_t which is attributed entirely to the improvement in technology. Foreign banks showed M_t growing, while domestic banks record a slight decline although the difference is not substantial. Both groups had an improvement in technology and for domestic banks the improvement is more apparent in the scale efficiency. Regarding the change in M_t by bank size, the results are not so clear. For example, all size groups recorded an improvement in technology, while larger banks (e.g. in third and fourth quartile) made an improvement in ΔPTE but faced a decline in ΔSE.

Table 9. Malmquist Total Factor Productivity Change Index, (2002-2005)

|       | ΔPTE | ΔSE | ΔTE | ΔT | M_t |
|-------|------|-----|-----|----|-----|
| Average | 0.978 | 0.999 | 0.977 | 1.032 | 1.008 |
| Year   |      |     |     |    |     |
| 2003 vs. 2002 | 0.941 | 0.972 | 0.914 | 1.141 | 1.043 |
| 2004 vs. 2003 | 1.017 | 1.032 | 1.048 | 0.950 | 0.998 |
| 2005 vs. 2004 | 0.976 | 0.994 | 0.968 | 1.009 | 0.980 |
| Country |      |     |     |    |     |
| Bulgaria | 0.979 | 0.998 | 0.977 | 1.035 | 1.011 |
| Croatia | 0.967 | 0.992 | 0.959 | 1.030 | 0.988 |
| Kosova | 1.022 | 1.050 | 1.064 | 1.029 | 1.096 |
| Ownership |      |     |     |    |     |
| Foreign banks | 0.997 | 0.985 | 0.981 | 1.036 | 1.016 |
| Domestic banks | 0.955 | 1.019 | 0.972 | 1.028 | 0.998 |
| Size |      |     |     |    |     |
| First Quartile | 0.962 | 1.009 | 0.970 | 1.019 | 0.988 |
| Second Quartile | 0.981 | 0.999 | 0.978 | 1.042 | 1.020 |
| Third Quartile | 1.004 | 0.993 | 0.997 | 1.044 | 1.040 |
| Fourth Quartile | 1.004 | 0.967 | 0.971 | 1.045 | 1.015 |

Note: ΔPTE is the technical efficiency change due to VRS assumption; ΔSE is the scale efficiency change; ΔTE is the technical efficiency change due to the CRS assumption; ΔT is the technological change and M_t is the total factor productivity change.

In sum, a minor improvement is apparent in M_t in the SEE banking sector, driven mainly by the change in technology. The largest improvement in all the M_t components is apparent for Kosovo in the period under review and foreign banks show some improvement in M_t as compared to their domestic counterparts. A shift in technology may be attributed to the fast technology advancement in Kosovo primarily because the banking sector had to introduce the whole set of banking products and services starting from branching network, ATMs, POS, e-banking, etc., which in turn were reflected in transitory cost inefficiencies. Other countries in the region were not subjected to the introduction
of the banking products and services from scratch and, hence, exhibited slower technology advancement. Overall, it seems that few banks took advantage of introducing new technologies while the average bank failed to keep up, leading to the divergence in the sector.

8. Conclusions

In this paper the intermediation efficiency of banks in four SEE countries over the period 2002-2005 was explored employing non-parametric techniques. The intermediation efficiency was measured by utilizing a two input-two output model for banks under investigation. Several findings emerged from the analysis. First, in the period 2002-2005, the intermediation efficiency of banks in the region as a whole only showed a marginal improvement. This is particularly because banks, on average, failed to improve their efficiency as much as the few best performers that exhibited a shift in technology, giving a continuously widening divergence. Inefficiencies in the sector emanated from substantial scale and cost inefficiencies, lending shortfalls and possibly the regulatory measures imposed by the central banks in the region aimed at curbing the rapid credit expansion. Cost and scale inefficiencies are partly a reflection of the technology shift effect, i.e. due to transitory costs in the adoption of technological advances for upgrading banking products and services, a shift towards the segments of business such as the small business sector which are more costly to serve and partly due to mergers and acquisitions (M&As) that led to an inefficient scale of operation.

Second, it was found that foreign banks (which are on average larger) were more efficient than domestic players and the gap had continuously widened. Foreign banks’ superior efficiency is identified despite the consolidation process through M&As largely involving these banks as in Croatia; start-up costs or costs of acquiring domestic banks; costs adopting new technologies or building them from scratch as in Kosovo; and costs in further building human capital which was generally lacking in the SEE. Third, in most cases larger banks were found to be more efficient than small banks, although experiencing substantial scale inefficiencies and exhibiting decreasing returns to scale, i.e. larger banks operated at excessive size. Fourth, regarding individual countries, the Kosovo banking system was found to be the least efficient in intermediation compared to other countries, although improvement and a catching-up effect has taken place over time. This is particularly because of the young banking sector that needed to expand its activities reflected in substantial scale and cost inefficiencies in the short-run and the risk perception of bankers and/or the regulator entailing lending shortfalls and high provisioning.

Some important policy guidelines emerge from the analysis. Given scale inefficiencies detected in the banking sectors in the SEE in the period under review, especially for larger/foreign banks exhibiting decreasing returns to scale, policymakers should carefully examine any M&A proposal involving the large players. First, M&As should be assessed on the basis of whether they create efficiency gains or increase the market power in already concentrated markets. A common motive for consolidation is gains in scale and scope economies, cost saving and revenue enhancement. Evidence from this paper suggests that these aims may have not been fulfilled, at least in the short-run. The absence of efficiency improvement due to M&As may indicate some degree of market power exploitation. Although there may be positive benefits of consolidation over a longer period than considered in this study (with changing policies and procedures and staff training to adaptation to new conditions, staff turnover, changes in customer relationships, change in organizational cultures, etc.), the evidence presented here is an important factor for policy makers to take account of in decision making. In terms of efficiency improvements, it is possible that M&As involving small/domestic banks that generally exhibit increasing returns to scale would be more appropriate since the DEA results, in line with most of other findings in the literature, suggest that the scale efficiencies in banking are exhausted early.

Second, in so far as M&As lead to a small number of large players in the market, policy makers may be concerned that the new large banks may be oriented towards lending to large corporate clients (e.g. Berger et al., 2001 and DeHaas et al., 2007 for recent evidence in TEs). Hence, the small business sector may be left relatively underserved and credit constrained. To the extent that large banks have a competitive disadvantage in lending to small firms, several lending technologies, other than those based on financial statement data which may be unavailable to small firms, has been proposed in order to overcome this disadvantage (Berger and Udell, 2006): collateral-based lending, credit-scoring, factoring, leasing, etc. However, institutional quality plays a crucial role in, among others, the ability of banks to adopt various lending technologies. For example, collateral-based lending and factoring depends directly on the enforcement of collateral and other laws on contracts and lending based on credit-scoring depends directly on the functioning of credit registers and other information infrastructures. Hence, poor institutions make it ambiguous whether these lending technologies can be advanced to overcome a competitive disadvantage of large banks to lending.
to small firms. In the view of competitive pressures from large/often foreign banks, on small/often domestic banks, the latter need to find market niches such as targeting small firms. Small domestic banks may have a competitive advantage serving the small business sector based on relationship lending and 'soft' information.

The central banks should pay careful attention to banks that fail to acquire new technologies and become increasingly inefficient as they may be forced to exit the market and create instability in the banking sector. This is particularly important as the detected inefficiencies in the literature using DEA have been found to have a good predictive power for bank failures. As shown in this paper, the sensitivity tests indicate that the DEA efficiency scores are stable over time, not sensitive to outliers and broadly consistent with ratio analysis. One implication for the central banks is that, apart from methods utilized in assessing bank performance such as the CAMEL technique (which is based on financial ratios), the adoption of DEA could be a supplementary early warning tool. However, this depends crucially on the appropriate statistical information and supervisory capacity in detecting misreporting on time. Also one possible explanation for the intermediation inefficiencies detected using DEA was the regulatory measures imposed by the central banks to curtail credit expansion, such as supplementary reserve requirements for banks in which credit growth is higher than the growth in total assets. Given the predominance of household loans in bank portfolios, these measures may disproportionately affect credit to the enterprise sector and subsequently small firms. Hence, an implication for the central banks is that the imposed measures targeting the restrain of credit growth should discriminate between household and lending to enterprises.

Acknowledgements

An earlier version of this paper appeared as CBK Working Paper No.4 and presented at the First Research Conference “Activating Sources of Economic Growth in Kosovo”, Riinvest: Prishtina, 14 May 2010. The authors would like to thank Professor Jean Mangan for useful comments.

References

Aly, H., R. Grabowski, C. Pasurka & N. Rangan (1990). Technical, Scale, and Allocative Efficiencies in U.S. Banking: An Empirical Investigation. Review of Economics and Statistics, 72, 211 - 218.

Amel, D., C. Barnes, F. Panetta & C. Salleo (2004). Consolidation and Efficiency in the Financial Sector: A Review of the International Evidence. Journal of Banking and Finance, 28, 2493 - 2519.

Banker, R., D. Charnes & A. Cooper (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science, 30, 1078 - 1092.

Barr, R., L. Seiford & T. Siems (1994). Forecasting Bank Failure: A Non-Parametric Frontier Estimation Approach. Recherches Economiques de Louvain, 60, 411 - 429.

Bauer, P., A. Berger, G. Ferrier & D. Humphrey (1998). Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. Journal of Economics and Business, 50, 85 - 114.

Berger, A. & L. Mester (2003). Explaining the Dramatic Changes in Performance of U.S. Banks: Technological Change, Deregulation, and Dynamic Changes in Competition. Journal of Financial Intermediation, 12, 57 - 95.

Berger, A., L. Klapper & G. Udell (2001). The Ability of Banks to Lend to Informationally Opaque Small Businesses. Journal of Banking and Finance, 25, 2127 - 2167.

Berger, A. & G. Udell (2006). A More Complete Conceptual Framework for SME Finance. Journal of Banking and Finance, 30, 2945 - 2966.

Bonin, J., I. Hasan & P. Wachtel (2004). Banks Performance, Efficiency and Ownership in Transition Countries. Journal of Banking and Finance, 29, 31 - 53.

Bos, J. & C. Kool (2006). Bank Efficiency: The Role of Bank Strategy and Local Market Conditions. Journal of Banking and Finance, 30, 1953 - 1974.

Caves, D., L. Christensen & W. Diewert (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity. Econometrica, 50, 1393 - 1414.
Charnes, A., W. Cooper & E. Rhodes (1978). Measuring the Efficiency of Decision Making Units. European Journal of Operational Research, 2, 429 - 444.

Coelli, T. (1996). A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program. CEPA Working Paper, No. 96/08.

Colwell, R. & E. Davis (1992). Output and Productivity in Banking. Scandinavian Journal of Economics, 94, 111 - 129.

Cooper, W., L. Seiford & K. Tone (2000). Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, Boston MA: Kluwer Academic Publishers.

DeHaas, R., D. Ferreira & A. Taci (2007). What Determines Banks’ Customer Choice? Evidence from Transition Countries. EBRD Working Paper, No. 104.

Färe, R., S. Grosskopf & C. Lovell (1994). Production Frontiers, Cambridge MA: The Cambridge University Press.

Ferrier, G. & K. Lovell (1990). Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence. Journal of Econometrics, 46, 229 - 245.

Fiorentino, E., A. Karmann & M. Koetter (2006). The Cost Efficiency of German Banks: A Comparison of SFA and DEA. Deutsche Bundesbank Discussion Paper, No. 10/2006.

Fries, S. & A. Taci (2004). Cost Efficiency of Banks in Transition: Evidence from 289 Banks in 15 Post-Communist Countries. EBRD Working Paper, No. 86.

Green, C., J. Murinde & I. Nikolov (2003). Are Foreign Banks in Eastern Europe More Efficient Than Domestic Banks? presented at 24th SUERF Colloquium on Stability and Efficiency Of Financial Markets in Central and Eastern Europe, Tallinn, Estonia, June 12-14, 2003.

Grifell-Tatjé, E. & C. Lovell (1997). The Sources of Productivity Change in Spanish Banking. European Journal of Operational Research, 98, 364 - 380.

Grifell-Tatjé, E. & C. Lovell (1999). A Note on the Malmquist Productivity Index. Economic Letters, 47, 169 - 175.

Grigorian, D. & V. Manole (2002). Determinants of Commercial Bank Performance in Transition: An Application of Data Envelopment Analysis. World Bank Working Paper, No. 2850.

Hasan, I. & K. Marton (2003). Development and Efficiency of the Banking Sector in a Transitional Economy. Journal of Banking and Finance, 27, 2249 - 2271.

Havrylchyk, O. (2006). Efficiency of the Polish Banking Industry: Foreign versus Domestic Banks. Journal of Banking and Finance, 30, 1975 - 1996.

Isik, I. & M. Hassan (2002). Technical, Scale and Allocative Efficiencies of Turkish Banking Industry. Journal of Banking and Finance, 26, 719 - 766.

Isik, I. & M. Hassan (2003). Financial Deregulation and Total Factor Productivity Change: An Empirical Study of Turkish Commercial Banks. Journal of Banking and Finance, 27, 1455 - 1485.

Jemric, I. & B. Vujicic (2002). Efficiency of Banks in Croatia: A DEA Approach. Comparative Economic Studies, 44, 169 - 193.

Kao, C. & S. Liu (2004). Predicting Bank Performance with Financial Forecasts: A Case of Taiwan Commercial Banks. Journal of Banking and Finance, 28, 2353 - 2368.

Kasman, A. & C. Yildirim (2006). Cost and Profit Efficiencies in Transition Banking: The Case of New EU Members. Applied Economics, 38, 1079 - 1090.

Kraft, E. & D. Tirtiroglu (1998). Bank Efficiency in Croatia: A Stochastic-Frontier Analysis. Journal of Comparative Economics, 26, 282 - 300.

Kraft, E., R. Hofler & J. Payne (2002). Privatization, Foreign Bank Entry and Bank Efficiency in Croatia: A Fourier-Flexible Function Stochastic Cost Frontier Analysis. Croatian National Bank Working Paper, No. 9.
Leong, W., B. Dollery & T. Coelli (2003). Measuring the Technical Efficiency of Banks in Singapore for the Period 1993-1999. *ASEAN Economic Bulletin*, 20, 195 - 210.

Matousek, R. (2004). Efficiency and Scale Economies in Banking: Empirical Evidence from Eight New EU Countries. Centre for International Capital Markets, London Metropolitan University Discussion Paper, No. 04-6.

McAllister, P. & D. McManus (1993). Resolving the Scale Efficiency Puzzle in Banking. *Journal of Banking and Finance*, 17, 389 - 405.

Mester, L. (1996). A Study of Bank Efficiency Taking into Account Risk-Preferences. *Journal of Banking and Finance*, 20, 1025 - 1045.

Mlima, A. & L. Hjalmarsson (2002). Measurement of Inputs and Outputs in the Banking Industry. *Tanzanet Journal*, 3, 12 - 22.

Nikiel, E. & T. Opiela (2002). Customer Type and Bank Efficiency in Poland: Implications for Emerging Banking Market. *Contemporary Economic Policy*, 20, 255 - 271.

Pastor, J., F. Perez & J. Quesada (1997). Efficiency Analysis in Banking Firms: An International Comparison. *European Journal of Operational Research*, 98, 395 - 407.

Ramanathan, R. (2003). *An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement*. New Delhi: Sage Publications.

Sabi, M. (1996). Comparative Analysis of Foreign and Domestic Bank Operations in Hungary. *Journal of Comparative Economics*, 22, 179 - 188.

Sealey, C. & J. Lindley (1977). Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions. *Journal of Finance*, 32, 1251 - 1265.

Simar, L. (2003). Detecting Outliers in Frontier Models: A Simple Approach. *Journal of Productivity Analysis*, 20, 391 - 424.

Stavarek, D. (2003). Banking Efficiency in Visegrad Countries Before Joining the European Union, presented at Workshop on Efficiency of Financial Institutions and European Integration, Technical University, Lisbon, Portugal, October 30-31, 2003.

Weill, L. (2003). Banking Efficiency in Transition Economies: The Role of Foreign Ownership. *Economics of Transition*, 11, 569 - 592.

Wheelock, D. & P. Wilson (1999). Technical Progress, Inefficiency, and Productivity Change in U.S. Banking, 1984-1993. *Journal of Money, Credit, and Banking*, 31, 212 - 234.

Yeh, Q. (1996). The Application of Data Envelopment Analysis in Conjunction with Financial Ratios for Bank Performance Evaluation. *Journal of the Operational Research Society*, 47, 980 - 988.

Yildirim, H. & G. Philippatos (2003). Efficiency of Banks: Recent Evidence from the Transition Economies of Europe, 1993-2000. *European Journal of Finance*, 13, 123 - 143.