EFFICIENCY EVALUATION OF PRODUCTION UNITS THROUGH DATA ENVELOPMENT ANALYSIS IN THE BANKING SECTOR

Fields of study: Systems Engineering and Informatics

Mgr. Ing. Lucie Chytilová

Ostrava, July, 2018
EFFICIENCY EVALUATION OF PRODUCTION UNITS THROUGH DATA ENVELOPMENT ANALYSIS IN THE BANKING SECTOR

Study program: Systems Engineering and Informatics
Fields of study: Systems Engineering and Informatics

Ostrava, July, 2018
Author of doctoral dissertation: Mgr. Ing. Lucie Chytilová
Department of Systems Engineering
Faculty of Economics
VŠB – Technical University of Ostrava

Department of Systems Engineering
Faculty of Economics
VŠB – Technical University of Ostrava

The defense of the doctoral dissertation will be held before a committee for the defense of doctoral dissertation in the field of study Systems Engineering and Informatics at the Faculty of Economics, VŠB – TU Ostrava, Sokolská tř. 33.

The doctoral dissertation can be consulted at the Department of doctoral studies, Havlíčkovo nábřeží 38a, room E 410.
# Table of contents

1. Introduction ......................................................................................................................... 2
2. Objective and structure of doctoral dissertation ................................................................. 4
3. Content of doctoral dissertation .......................................................................................... 5
4. Procedure of doctoral dissertation ...................................................................................... 7
5. Methods applied in doctoral dissertation ............................................................................ 9
6. Summary of results and conclusion ................................................................................... 20
7. List of references ............................................................................................................... 22
8. List of author’s publications and research ........................................................................ 35
9. Summary ........................................................................................................................... 39
10. Shrnutí .............................................................................................................................. 40
1 Introduction

An integral part of today's economic system is the financial sector. The financial sector has an extreme impact on macroeconomics and microeconomics. It affects the allocation of financial resources. Banking and other financial institutions play a significant role. In order to properly allocate the economic sources, they have to be efficient. The current global interconnection of all participants in the financial sector has many advantages - easier mobility of capital, more efficient allocation of resources, economies of scale in big companies, lower prices for consumers in countries with great economic progress and so on. On the other hand, the globalization is connected to certain disadvantages as well. If the economy has a problem, the impacts of the problems may have a great effect on each individual participant and on the entire economic system. This is due to the integration of financial markets and the liberalization of capital flow.

Monitoring, analysing and understanding to the situation in the financial sector should lead to a minimization of the problems. More precisely, efficiency measurements of the banking sector determine the performance of all evaluated units (banks) and help them to detect the possibility to improve. These measurements provide valuable knowledge to the market regulators and give them material for their decisions. The responsible institutions are able to set better rules for management and supervision in the financial sector if they know the real state of the efficiency score compared to others.

This doctoral dissertation focuses on the creation of new model/models of the Data Envelopment Analysis (DEA), which will serve to better monitor, analyse and understand to the situation in financial sector of the Visegrad Group (VG) countries. In other words, a system of new models is developed for this current problem area of financial institutions in the VG, which should lead to an assessment of efficiency, the detection of inefficiencies and the possibility of elimination.

After the revolution in the late 80s and 90s, the VG countries were forced to face many problems. The main problem in the field of economy was the long-term impact of the centrally planned economy with a minimal role of the market. This type of economy lasted over 45 years in these countries. Consequences of this policy had many forms, for example: inherited bad loans, the lack of experience in commercial banking, the lack of skilled workers in the banking sector, rapidly growing number of banks, privatization of state-owned banks, entry of foreign owners, changes in legislation, emergence of supervisors and supervision, etc.

The 90s saw in a sign of growth for the VG countries. The bank market has increased its competition due to the entry of foreign banks into domestic markets and the emergence of new banking products. Unfortunately, indebtedness of population and companies had grown at a robust pace. This resulted in a very sensitive banking system. In the context of the sensitive banking system, some negative economic factors have been seen - supply and demand shocks. Since 2000, all the VG countries have been referred to as the countries with a stabilized banking sector and large interconnection to foreign banks. In May 1, 2004, all the VG countries became the European Union (EU) member states.

\[^{1}\text{scale of EBRD - European Bank for Reconstruction and Development}\]
The foreign and political activities of the VG have significantly increased after accession to the EU. The VG countries have focused on promoting cooperation and stability in the wider region of Central Europe. In addition, the economic situation started to improve, until the economic crisis. The Central and Eastern Europe (CEE) countries, including the region of the VG countries, were one of the most attractive regions for foreign investors before the financial crisis. The share of foreign investors in the banking sector was 80% on average (Miklaszewksa et al. 2012). However, this led to a large expansion of the loans, which caused great sensitivity of banking systems to the financial crisis and the development of these countries slowed down rapidly. The VG countries were preparing to adopt the single currency of the EU, the Euro. So far, only Slovakia has joined the third stage of the European Monetary Union (in 2009).

The possibility to analyse a group of countries like this, is in their similarities - similar development in the banking system, same current structure of the banking system, interconnectedness of the banking market in the form of the same parent banks, size of the banking sector, strong capital representative offices of foreign banks, EU membership (similar legislation and regulation) and so on.

The similarities given above are obvious. However, there should still be some differences. The VG countries have undergone many changes in recent years - they had to build new commercial banking, they had to join the EU and face the financial crisis. Probably none of these things happened in the same exact time and way in all VG countries. Therefore, if the banks' efficiency is calculated, analysed and closely compared, some differences among the VG counties should be expected. Overall, the development of the past may give important information for future.
2 Objective and structure of doctoral dissertation

The main goal of the doctoral dissertation is to create and modify the basic Data Envelopment Analysis (DEA) models and apply these new models to evaluate the financial sector in the VG countries during the time period 2005-2015.

The following partial scientific questions lead to the sub-goals:

- to select the most appropriate DEA model for the given area with the assumption of missing data (DEA model with interval data);
- to modify the best selected DEA model with use of a dual-role variable to better understand the situation in the given area;
- to extend the modified DEA model with a risk variable model, which should help to better describe the situation in the given area;
- to compare and prove that productivity changes based on an alternative total production factor (Hicks) is better than a classical total production factor (Malmquist) (all based on the modified models).

The doctoral dissertation has the following structure. The following chapter, Chapter 2, is a theoretical part of the doctoral dissertation. Information is introduced about general production as well as the specific approaches in production of financial institutions. Chapter 3 provides information about the current state of the financial sectors in the VG countries. In Chapter 4 are described all applied methods and approaches. There can be found basic DEA models as well as all the modification of them. Also, the description of the alternative productivity indices can be found. Chapter 5 deals with all applications of the DEA models. In the end of the doctoral dissertation is also the bibliography and annexes with all the results as lists of all models, tables and pictures.

---

2 The doctoral dissertation had to deal only with commercial banks in the VG, but due to the data deficiencies and assumptions of DEA methods, other types of banks have been used as well. From this point on, instead of "banking sector", the term "financial sector" will be used.

3 There had not been more data for the analysis, when this doctoral dissertation was started. The problem of the financial institution environment is that financial institutions do not like to provide information. If the information is provided, the proper and reliable information is given with at least 8 months' delay. It means that in the fall of 2017 data would be available for the year 2016, but at this time, this doctoral dissertation thesis was almost all done.

4 In this doctoral dissertation, the missing data are represented by a defined interval.
3 Content of doctoral dissertation

1 Introduction ................................................................. 8
   1.1 Motivation of doctoral dissertation ........................................ 8
   1.2 Research questions of doctoral dissertation ............................ 10
   1.3 Organization of doctoral dissertation ..................................... 10

2 Theoretical background of financial institution production ................. 12
   2.1 Production possibilities of financial institutions .......................... 12
   2.2 Qualification of production variables for production of financial institutions ........................................................................ 14
      2.2.1 Production approach ................................................................. 15
      2.2.2 Intermediation approach .............................................................. 16
      2.2.3 Profit-oriented approach .............................................................. 16
   2.3 Literature overview of current state in financial industry ............... 17
      2.3.1 DEA and SFA ........................................................................ 18
      2.3.2 DEA around the world ............................................................... 19
      2.3.3 DEA in Visegrad Group countries ............................................. 20
   2.4 Summary .................................................................................. 28

3 Evaluation of financial systems in the Visegrad Group countries ............. 29
   3.1 The Czech Republic .................................................................. 29
   3.2 Hungary .................................................................................... 36
   3.3 Poland ....................................................................................... 43
   3.4 Slovakia ..................................................................................... 50
   3.5 Summary .................................................................................. 57

4 Description of applied methods and approaches ..................................... 59
   4.1 Division of used methods ............................................................ 59
   4.2 Data Envelopment Analysis ....................................................... 60
      4.2.1 Non-allocation DEA models ...................................................... 63
      4.2.2 Allocation DEA models ............................................................ 67
      4.2.3 DEA for missing data ............................................................... 68
      4.2.4 DEA with dual-role variable ..................................................... 70
      4.2.5 DEA with risk variable ............................................................. 73
      4.2.6 Over all model ........................................................................ 74
   4.3 Total factor productivity indices .................................................. 76
      4.3.1 Measures of efficiency by TFP .................................................. 77
      4.3.2 The components of TFP change ............................................... 82
4 Procedure of doctoral dissertation

The doctoral dissertation can be divided into four logical parts:

- theoretical background of financial institutions production;
- evaluation of financial systems in the Visegrad Group countries;
- description of applied methods and approaches;
- analysis and evaluation of banking efficiency in Visegrad Group countries.

At the beginning, it is necessary to emphasize the importance of financial institutions and its production. The doctoral dissertation solves efficiency of financial institutions based on the understanding of their production (outputs). Therefore, it is important to understand the theoretical fundamentals of financial institutions’ production. This means that it is important to define the production capabilities of financial institutions and how these capacities relate to estimated production functions. In addition, the information about the current examination in the field of financial institutions is shown.

The following part deals with a description and evaluation of financial systems in each country of the VG. The information about the current state of the financial sectors of each country helps to better understand the results of analysis and check if the results are in logical link with the economic development in the countries.

The third part deals with a description of quantitative methods. The theoretical basis behind the used methodology and approaches is explained. First, only a brief introduction into the used methodology is given, then the basic DEA models are introduced. More precisely, there are introduced two basic models - model by Charnes, Cooper, and Rhodes so called the CCR model and model by Baker, Charnes, and Cooper model so called the BCC input oriented models. The ability of allocation models to calculate the cost, technical, and allocative efficiencies is also described. Then the more complicated DEA models are introduced - models for missing data or the dual-role variable. Finally, two new models, based on the previous information, are defined. The above-mentioned modelling process is shown in Figure 1 as well as the scientific questions. The end of this part deals with total factor productivity indices. The connection between efficiency and total factor productivity (TFP) is given as well as the opportunities to divide the classical efficiency into more parts - technical efficiency, scale efficiency, mix efficiency, and so on.

The quantitative methods are further applied to the data set of financial industry in Visegrad Group countries at the time period from 2005 to 2015. The calculations are following the scientific questions and the quantitative methods. First, the basic DEA models are calculated then the more complicated and finally the new models are used. As well as the Malmquist index and Hicks index. All the calculation is done in the software GAMS (all codes are done by author). The final output of the calculation provides new and deeper information about the financial institutions in the VG. With these information managers of financial institutions are able to manage these institutions better.
Figure 1: The structure of DEA model in accordance with initial goals and conclusion

Source: author’s own processing, 2017
5 Methods applied in doctoral dissertation

The basic theme of this doctoral dissertation is to measure efficiency and productivity changes. This doctoral dissertation uses the non-parametric method - the Data Envelopment Analysis (DEA). This is based on previous analyses, which have detected that one of the most often used method in the area of banking industry is the DEA method.

Data Envelopment Analysis

The Data Envelopment Analysis is the non-parametric approach to measure efficiency. It is based on mathematical programming. This approach was first formulated by Farrell (1957) as a single-input/output model with radial measure of technical efficiency without the requirement of any price information. Farrell introduced concepts of the productivity and the efficiency in his work. Farrell's model was generalized into the multiple input/output case and reformulated as a mathematical problem by Charnes et al. (1978).

The further development of DEA was very intense and was influenced by many researchers, e.g., Handbook on Data Envelopment Analysis by Cooper et al. (2004) and their earlier publications. The great interest in DEA method might be explained by the fact that this method has many advantages and it is well applicable. For example, DEA method may use multiple input and output variables without any aggregation. This is suitable in applications where the price information is not available or is hardly accessible, e.g., banking industry, health care. DEA methods (as the non-parametric method) assumes that the production function does not to be specified. The basic and important assumption for DEA method is the assumption of the convexity - distinctive Free Disposable Hull (FDH) which envelops the data.

The principle of DEA use might be simply described in two steps. First, the efficiency frontier has to be computed. Second, the efficiency scores of all DMUs have to be obtained and identified. Then the inefficient DMUs could be compared with efficient counterparts. These inefficient DMUs have to lie below the efficient frontier. Before the efficiency frontier is computed, the decision on the assumption of production has to be made (based on theory for the field of use) - Constant Return to Scale (CRS) or Variable Return to Scale (VRS).

The classic DEA method determines the score of technical efficiency (TE). DMUs on the efficiency frontier (efficient DMUs) may be known as the Pareto-efficient DMUs. The efficiency may be measured as input or output oriented. The decision depends on the purpose of the study or the general idea of efficiency evaluation.

The input oriented model represents a model where the Pareto-efficient DMUs use a minimum possible amount of controllable productive input sources to generate outputs. The inefficient DMUs may be moved to the efficiency frontier by reducing the consumption of certain input or inputs. In the output oriented model, DMUs have to produce the maximum amount of controllable outputs by a given amount of inputs. In this case, the inefficient DMUs may be shifted to the efficient frontier by increasing the certain output or outputs. All Pareto-efficient DMUs have a benchmark efficiency score equal to unity. These DMUs are not able to improve any input or output without deterioration of other input or output. If the DMUs are not Pareto-efficient, the efficiency score is in the interval \([0, 1)\). The last type of models are non-oriented (additive) models. These models are based on the optimal mix of inputs and outputs.
The one criterion for the classification of DEA approaches is the possible orientation, as it is mentioned above. In input oriented DEA model, the technical efficiency of production unit is measured as follows:

\[
\theta_{\text{input}} = \frac{\text{minimum input}}{\text{actual input}}
\]  

(1)

Another criterion for the DEA models may be distribution into allocation and non-allocation DEA models. The difference of these two models is in the use of the information about the price of inputs/outputs. The non-allocation DEA models determine the technical efficiency of each DMUs without using any information on prices. The technical efficiency evaluates the physical transformation of production inputs to outputs relative to other DMUs with the use of the certain technologies. The allocative models, on the other hand, needs to know the prices. Both types of models are described closely in the following subsections.

Non-allocation DEA models

The non-allocation DEA models should be divided into:

- CCR model presented in the work by Charnes et al. (1978);
- BCC model presented in the work by Banker et al. (1984).

The CCR DEA model by Charnes et al. (1978) was developed based on the assumption of the constant returns to scale. The CCR model assumes a strong disposability of inputs and outputs and also the convexity of the production possibility set. The application of the CCR model provides the technical efficiency score for each DMU. It also gives information about input and output slacks and a reference set for the efficient DMUs. The doctoral dissertation is, inter alia, dealing with the input oriented CCR model. Note, this model is referred as a CCR-I model.

To define a multiplier form of CCR-I model in linear form, let assume \( n \) DMUs \((j = 1, 2, ..., n)\). Each DMU is producing \( y \in R^s_+ \) different quantity of outputs by using \( x \in R^m_+ \) different quantity of inputs. Variable \( x_{ij} \) is denoted as the amount of the \( i^{th} \) input used by the DMU \( j \) \((i=1, 2, ..., m)\). Variable \( y_{rj} \) expresses the amount of the \( r^{th} \) output produced by the DMU \( j \) \((r=1, 2, ..., s)\). A linear combination of multiple inputs and multiple outputs for each DMU allows to obtain the technical efficiency for the target 'o' by solving the following fractional programming model:

\[
\max_{\mu} f_o(\mu) = \sum_{r=1}^{s} \mu_{r} y_{r o}
\]

s.t.

\[
\begin{align*}
\sum_{i=1}^{m} v_i x_{i o} &= 1 \\
\sum_{r=1}^{s} \mu_{r} y_{r j} - \sum_{i=1}^{m} v_i x_{ij} &\leq 0, \quad j = 1, ..., n \\
\mu_{r} &\geq \epsilon, \quad r = 1, ..., s \\
v_i &\geq \epsilon, \quad i = 1, ..., m.
\end{align*}
\]

(2)

where \( y_{r o} \) is the amount of the \( r^{th} \) output produced by the DMU \( o \), \( x_{i o} \) is the amount of the \( i^{th} \) input used by the DMU \( o \), \( \mu_{r} \) is the weight given to output \( r \), \( v_i \) is the weight given to input \( i \) and \( \epsilon \) is a non-Archimedean (infinitesimal) constant \((\epsilon > 0)\).
The results obtained by solving the optimization problem (2) provide weights for all inputs and outputs that maximize the proportion of a virtual output to a virtual input for DMU\(_o\). The maximum efficiency score \(f_0^*\) is equal to 1. Since the number of limitation for the primary model (2) is very high \((n+s+m+1)\). Charnes et al. (1978) also created a dual program to the primary model (2). This dual programme is in envelopment form, see Charnes et al. (1978).

The BCC DEA by Banker et al. (1984) has been developed as an extension of the CCR model. The difference is in the production assumption. The BCC model is defined with the assumption of variable returns to scale. The advantage of the BCC model is that the technical efficiency of this model may be divided into the pure technical efficiency (PTE) and the scale efficiency (SE). Note, the input oriented BCC model is referred to as BCC-I model. The BCC-I model measures the pure technical efficiency of the DMU\(_o\) by solving the primary optimization problem in the multiplier form as follows:

\[
\begin{aligned}
\max_{\mu, \mu_o} f_0 (\mu, \mu_o) &= \sum_{r=1}^s \mu_r y_{ro} - \mu_o \\
\text{s.t.} \\
\sum_{i=1}^m v_i x_{io} &= 1 \\
\sum_{r=1}^s \mu_r y_{rj} - \sum_{i=1}^m v_i x_{ij} - \mu_o &\leq 0, \quad j = 1, \ldots, n \\
\mu_r &\geq \epsilon, \quad r = 1, \ldots, s \\
v_i &\geq \epsilon, \quad i = 1, \ldots, m \\
\mu_o &\text{free in sign.}
\end{aligned}
\]  

(3)

The multiplier models (2) and (3) differ just in one variable - the free variable \(\mu_o\) in the BCC-I model (3). The free variable is used in both types of orientation. There are also dual models which are introduced in this doctoral dissertation. The relationship between the optimal objective values of the CCR-I model and the BCC-I model is following:

\[
f_0^{*\text{BCC}} \geq f_0^{*\text{CCR}}.
\]  

(4)

The value of scale efficiency for DMU\(_o\) is given by the following relationship:

\[
\text{SE} = \frac{f_0^{*\text{CCR}}}{f_0^{*\text{BCC}}}.
\]  

(5)

The group of the non-allocation models includes the CCR and the BCC models with the CRS and the VRS assumptions, respectively. Also it includes the additive models, the multiplicative models and the slack-based measures (SBM) models (Kumar and Gulati, 2014). These models are not used in the doctoral dissertation, so they are just briefly mentioned below.

The additive models (Pareto-Koopmans models) are not based on the orientation. They provide a non-oriented measurements. These models simultaneously reduce inputs and augment outputs by taking the slacks. For example, Charnes et al. (1985) mentioned this group of models in their publications.

The multiplicative models were devoted by Charnes et al. (1982). These models measure the efficiency as the ratio of the weighted multiplicative product of outputs divided by the weighted multiplicative product of inputs in order to account for interdependencies between input or output.
The slack-based measures models were first presented by Tone (2001). Tone was examining inputs/outputs individually contrary to the radial approach that assumes the proportional changes in inputs/outputs. The scalar measure deals directly with the input excesses and the output shortfalls increasing in each input and output slack. The measure is determined only by its reference set. It means, it is not affected by statistics over the whole data set.

**Allocation DEA models**

The allocation DEA models are used to minimize cost and maximize revenue and profit in the same time. In as much as the doctoral dissertation is focused on the cost efficiency, the allocation DEA models are introduced also.

The calculation of the traditional cost efficiency (CE) assumes the knowledge of input market prices for each DMU. If the objective of the DMU is to minimize the cost, then the measure of the cost efficiency is provided by the following ratio:

$$ C_E_o = \frac{\text{minimum cost}}{\text{actual input}} = \frac{\sum_{i=1}^{m} p_o^i \hat{x}_{io}^*}{\sum_{i=1}^{m} p_o^i x_{io}} \tag{6} $$

where $p_o^i$ is the unit price of $i^{th}$ input for DMU$_o$, $\hat{x}_{io}^*$ is the optimal quantity of $i^{th}$ input for DMU$_o$ that minimizes the cost, $x_{io}$ is the actual value of the $i^{th}$ input for DMU$_o$. The ratio $C_E_o$ of (6) is situated in the interval between 0 and 1.

Farrell (1957) in his work estimated that the input oriented technical efficiency is just one component of the cost efficiency. The second component of the input oriented CE is the input oriented allocative efficiency (AE). AE reflects the ability of DMU to choose the inputs in optimal proportions given by their respective prices. Therefore, AE expresses whether the examined DMU uses the right mix of inputs due to their relative prices. AE is defined as a residual component of the cost efficiency of each DMU and is obtained as follows:

$$ AE = CE / TE \tag{7} $$

Note, in order to be cost efficient, the DMU must become technically efficient first.

The traditional cost efficiency by the linear programming problem is as follows:

$$ \min_{\lambda, \hat{x}} \sum_{i=1}^{m} p_o^i \hat{x}_{io} \tag{8} $$

s.t.  
$$ \sum_{j=1}^{n} \lambda_j x_{ij} \leq \hat{x}_{io}, \quad i = 1, \ldots, m $$
$$ \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \ldots, s $$
$$ \lambda_j, \hat{x}_{io} \geq 0, \quad j = 1, \ldots, n, $$

where $p_o^i$ is the unit price of $i^{th}$ input for DMU$_o$, $\hat{x}_{io}$ is the quantity of $i^{th}$ input for DMU$_o$ that minimizes the cost, $\hat{x}_{io}^*$ is the optimal value of $\hat{x}_{io}$, $x_{ij}$ is the actual value of the $i^{th}$ input for $j^{th}$ DMU. In the cost efficiency model, the unit cost for the DMU$_o$ has to be fixed at $p_o$ and the cost-minimizing input-bundle $\hat{x}^* = (x_{1o}^*, x_{2o}^*, \ldots, x_{mo}^*)$ that produces the output $y_{ro}$, is to be found.

Note, model (9) is the CRS input oriented cost efficiency model. The model (8) could be extended into the VRS case by adding the constraint: $\sum_{j=1}^{n} \lambda_j = 1$. 

12
DEA for missing data

There exist many improvements of the classical DEA models. To deal with the issue of missing data (interval data in this case) the special model has to be used. This special model is defined below. As an inspiration, the model of Smirlis et al. (2006) has been used. Their model was done for the output oriented DEA model, so the transformation was needed - input oriented DEA model with the VRS assumption.

Assume, there are \( n \) DMUs, each using \( m \) inputs to produce \( s \) outputs. For any unit \( j \) \((j = 1, 2, ..., n)\), the level of its \( r^{th} \) output \((r = 1, ..., s)\) is denoted by \( y_{rj} \) and by \( x_{ij} \) the level of its \( m^{th} \) input \((i = 1, ..., m)\). Unlike the original DEA model, the interval DEA assumes that some of the crisp input \( x_{ij} \) and output \( y_{rj} \) values are not known and for them, it is only known that they lie within bounded intervals, i.e. \( x_{ij} \in [x_{ij}^L, x_{ij}^U] \) and \( y_{rj} \in [y_{rj}^L, y_{rj}^U] \), with the upper and lower bounds of the intervals \( x_{ij}^L, x_{ij}^U, y_{rj}^L, y_{rj}^U \) to be strictly positive constants.

To be able to introduce the intervals instead of exact data into the model (3), some transformation should be done.

The values \( x_{ij} \) and \( y_{rj} \) are expressed in terms of new variables \( s_{ij} \) and \( t_{rj} \), respectively, to convert the non-linear model to a linear one. These new variables locate the level of inputs and outputs within the bounded intervals \( [x_{ij}^L, x_{ij}^U] \) and \( [y_{rj}^L, y_{rj}^U] \), respectively, as follows:

\[
\begin{align*}
    x_{ij} &= x_{ij}^L + s_{ij}(x_{ij}^U - x_{ij}^L), \quad i = 1, \ldots, m; \quad j = 1, \ldots, n \quad \text{with} \quad 0 \leq s_{ij} \leq 1, \\
    y_{rj} &= y_{rj}^L + t_{rj}(y_{rj}^U - y_{rj}^L), \quad r = 1, \ldots, s; \quad j = 1, \ldots, n \quad \text{with} \quad 0 \leq t_{rj} \leq 1.
\end{align*}
\]

Applying the above transformation to model (3), the following linear model is obtained:

\[
\begin{align*}
\max f_0 &= \sum_{r=1}^{s} \mu_r \left(y_{rj0}^L + t_{rj0}(y_{rj0}^U - y_{rj0}^L)\right) - \mu_o \\
\text{s.t.} \quad &\sum_{i=1}^{m} v_i \left(x_{ij0}^L + s_{ijo}(x_{ij0}^U - x_{ij0}^L)\right) = 1 \\
&\sum_{r=1}^{s} \mu_r \left(y_{rj}^L + t_{rj}(y_{rj}^U - y_{rj}^L)\right) - \sum_{i=1}^{m} v_i \left(x_{ij}^L + s_{ij}(x_{ij}^U - x_{ij}^L)\right) - \mu_o \leq 0, \quad j = 1, \ldots, n \\
&\mu_r \geq \varepsilon, \quad 0 \leq t_{rj} \leq 1, \quad r = 1, \ldots, s \\
&v_i \geq \varepsilon, \quad 0 \leq s_{ij} \leq 1, \quad i = 1, \ldots, m \\
&\mu_o \text{ free in sign.}
\end{align*}
\]

It can be seen that for inputs and outputs, there are new terms \( v_i s_{ij} \) and \( \mu_r t_{rj} \), respectively. These new terms may be replaced by new variables \( q_{ij} = v_i s_{ij} \) and \( p_{rj} = \mu_r t_{rj} \) which meet the needed conditions. The model (10) can be rewritten as follows:
\[
\max f_o = \sum_{r=1}^{s} \left( \mu_r y_{rjo}^L + p_{rjo} (y_{rjo}^U - y_{rjo}^L) \right) - \mu_o
\]

s.t. \[
\sum_{i=1}^{m} (v_i x_{ijo}^L + q_{ijo} (x_{ijo}^U - x_{ijo}^L)) = 1
\]
\[
\sum_{r=1}^{s} (\mu_r y_{rj}^L + p_{rj} (y_{rj}^U - y_{rj}^L)) - \sum_{i=1}^{m} (v_i x_{ij}^L + q_{ij} (x_{ij}^U - x_{ij}^L)) - \mu_o \leq 0,
\]
\[
p_{rj} - \mu_r \leq 0, \quad r = 1, \ldots, s
\]
\[
q_{ij} - v_i \leq 0, \quad i = 1, \ldots, m
\]
\[
v_i, \mu_r \geq \varepsilon, \quad \text{for all } i, r
\]
\[
q_{ij}, p_{rj} \geq 0, \quad \text{for all } i, r, j
\]
\[
\mu_o \text{ free in sign.}
\]

In model (10), variables under estimation are the weights \( \mu_r, v_i \) and the new variables \( q_{ij}, p_{rj} \) that denote the level of input and output values within the bounded intervals. For more details see Smirlis et al. (2006). This model may be done for the CRS assumption as well.

**DEA with dual-role variable**

Another improvement of the classical DEA models is to involve in the calculation not just the classical variables (inputs and outputs), but include also a so-called "dual-role variable". The idea of the dual-role variable is that this variable is treated as both the input variable and the output variable within the classical DEA framework. There have been many researches in this topic. For example, Beasley (1990, 1995) had studied the efficiency of university departments. This study had treated research funding on both sides of variables. Beasley had introduced the model, based on CCR input oriented (linearized) model that was probably the most intuitive for the issue at that time. The model is as follows:

\[
\max f_o = \sum_{r=1}^{s} \mu_r y_{ro} + \gamma \omega_o
\]

s.t. \[
\sum_{i=1}^{m} v_i x_{io} + \beta \omega_o = 1
\]
\[
\sum_{r=1}^{s} \mu_r y_{rj} + \gamma \omega_j - \sum_{i=1}^{m} v_i x_{ij} - \beta \omega_j \leq 0, \quad j = 1, \ldots, n
\]
\[
\mu_r \geq \varepsilon, \quad r = 1, \ldots, s
\]
\[
v_i \geq \varepsilon, \quad i = 1, \ldots, m
\]
\[
\gamma, \beta \geq 0
\]

where members \( j \) of a set of \( n \) DMUs are to be evaluated in terms of \( s \) output variables \( Y_j = (y_{rj})_{r=1}^{s}, i \) input variables \( X_j = (x_{ij})_{i=1}^{m} \), and the particular variable (for example, research income) which are held by each DMU in the amount \( \omega_j \), and serves as both an input variable and output variable.
Although the model (11) looked good, there were some major issues. One problem was in the absence of constraint on multipliers \( \{ \mu_r \} \) and \( \{ v_i \} \), each DMU is calculated as fully efficient (100%). If the model (11) is solved, all DMUs are seen efficient (see the proof in work by Cook et al. (2006)). The second problem is the logical/illogical structure of the input oriented structure of model (11). In the case, if DMU had efficiency score of \( f \), then all inputs, including \( \omega_o \), were reduced by \( 1 - f \). On the other hand, this factor was not included on the output side where \( \omega_o \) was not reduced. Thus, model (11) treated \( \omega_o \) differently on the input from on the output side.

The correction of model structure (11) was improved by many authors in many different ways. In this dissertation thesis, the idea from the work by Banker and Morey (1986) and Cook et al. (2006) had been taken. The idea was to treat the dual-role variable \( \omega_o \), differently. Specifically, since \( \omega_o \) served as the output variable, and was hence generally expected to remain at its current level in the input oriented setting, it should be treated as being a non-discretionary variable on the input side.

The DEA model based on research by Banker and Morey (1986) was used. Its linear form is as followed:

\[
\begin{align*}
\text{max } f_o & = \sum_{r=1}^{s} \mu_r y_{ro} + \gamma \omega_o - \beta \omega_o, \\
\text{s.t.} & \\
\sum_{i=1}^{m} v_i x_{io} & = 1, \\
\sum_{r=1}^{s} \mu_r y_{rj} + \gamma \omega_j - \sum_{i=1}^{m} v_i x_{ij} - \beta \omega_j & \leq 0 \quad j = 1, \ldots, n \\
\mu_r & \geq \varepsilon, \quad r = 1, \ldots, s \\
v_i & \geq \varepsilon, \quad i = 1, \ldots, m \\
\gamma, \beta & \geq 0,
\end{align*}
\]

(12)

where \( \omega_o \) can serve as input variable or output variable and \( \gamma \) and \( \beta \) are weights for the dual-role variable. The input variable was moved to output side, but with the opposite sign. The model (12) is at the optimum if \( \gamma^* - \beta^* = 0 \). The proof may be seen in work by Cook et al. (2006). According to this, it has been obvious that there have existed three possibilities in regard to the sign of \( \gamma^* - \beta^* \), where \( \gamma^* \) and \( \beta^* \) are the optimal values of the model (14); \( \gamma^* - \beta^* > 0 \), \( = 0 \), \( < 0 \). The sign of \( \gamma^* - \beta^* \) could have important implications when the dual-role variable has been a resource that could be allocated across the DMUs. The aspect of the sign of \( \gamma^* - \beta^* \) could be a comment based on the standard (no dual-role factors present) Variable Returns to Scale model of Banker et al. (1984). More precisely, the three cases should be examined in regard to the sign of \( \omega_o \), which identifies the VRS status the BCC. For more details, see the work by Cook et al. (2006). Generally, the cases as it follows:

- If \( \gamma^* - \beta^* < 0 \) then the dual-role variable is "behaving like the input variable", hence less of this variable would improve its performance of efficiency ratio (in case of the VRS model, the sign of \( \omega_o > 0 \));
- If \( \gamma^* - \beta^* > 0 \) then the dual-role variable is "behaving like the output variable", hence more of the variable is better, and would lead to an increase in efficiency (in case of the VRS model, the sign of \( \omega_o < 0 \));
- If \( \gamma^* - \beta^* = 0 \) then the dual-role variable is at an equilibrium or optimal level (in case of the VRS model, the sign of \( \omega_o = 0 \)).
All of this information should be calculated for the future use - the reallocation of a dual-role variable by looking individual DMUs or pairs.

**DEA with risk variable**

There are two views on how to treat the risk: exogenously or endogenously. The results with exogenous treatment (Ataullah et al., 2004; Chang and Chiu, 2006) showed that the efficiency level is significantly correlated with the risk indicators. The second treatment was in order to analyse banks’ efficiency (Chiu and Chen, 2008; Girardone et al., 2004). However, the majority of studies adopted the overdue loan ratio as the substitute variable for risks. Certainly, this does not reflect the characteristic of uncertainty that risks display. The risk is defined as the uncertainty and the degree of risk varies with the asset value fluctuation and the manager’s attitude toward the risk. The risk may bring profit or loss to the asset value. In this context the basic function of capital helps to cover the possible loss incurred by taking risks. There are many ways to include the risk component into the model. For example, there are studies which employ the Value at Risk (VaR) based on the New Basel Capital Accord (Basel II) (Kao and Liu, 2000; Chen et al., 2013) or there are studies which use so called demography risk, which conclude many variables - age, gender, income and so on (Cooper et al., 2014). According to the situation in which there is measured banking industries from four countries, the demography risk variable is used.

Different types of the DEA models have been seen which had treated as usual variables (input or output variables), but also special models have been created, for example see work by Kao and Liu (2000) and Chen et al. (2013) based on the slack-based measure of efficiency by Tone (2001). In this case, the model previous model with missing (interval data) and dual-role variables are used.

**Overall new models**

The new models, which combines the previous knowledge of basic DEA models and special DEA models (10) and (12), are introduced in this part of the section. There are new models and they bring a new contribution to the environment of missing data, risk variables and dual-role variables. Such models have not been previously developed, and these models are the biggest contribution of the doctoral dissertation.

Assume, there are \( n \) DMUs, each using \( m \) inputs to produce \( s \) outputs. For any unit \( j \) \((j=1, 2, ..., n)\), the level of its \( r^{th} \) output \((r = 1, ..., s)\) is denoted by \( y_{rj} \) and by \( x_{ij} \) the level of its \( i^{th} \) input \((i = 1, ..., m)\). Plus, there is the particular variable (for example, research income, deposits etc.) which is held by each DMU in the amount \( w_j \), and serves as both an input variable and output variable. Unlike the original DEA model, the interval DEA assumes that some of the crisp input \( x_{ij} \), output \( y_{rj} \) and dual-role \( w_j \) values are not known and for them, it is only known that they lie within bounded intervals, i.e. \( x_{ij} \in [x_{ij}^L, x_{ij}^U] \), \( y_{rj} \in [y_{rj}^L, y_{rj}^U] \) and \( w_j \in [w_j^L, w_j^U] \), with the upper and lower bounds of the intervals \( x_{ij}^L \), \( x_{ij}^U \), \( y_{rj}^L \), \( y_{rj}^U \), \( w_j^L \), \( w_j^U \) to be strictly positive constants.

To be able to introduce the intervals instead of exact data into the model (3), some transformation should be done.

The values \( x_{ij} \), \( y_{rj} \) and \( w_j \) are expressed in terms of new variables \( s_{ij} \), \( p_{rj} \) and \( b_j \), respectively, to convert the non-linear model to a linear one. These new variables locate the level of inputs and outputs within the bounded intervals \([x_{ij}^L, x_{ij}^U]\), \([y_{rj}^L, y_{rj}^U]\) and \([w_j^L, w_j^U]\), respectively, as follows:
\[ x_{ij} = x_{ij}^l + s_{ij}(x_{ij}^U - x_{ij}^l), i = 1, ..., m; j = 1, ..., n \text{ with } 0 \leq s_{ij} \leq 1, \]
\[ y_{rj} = y_{rj}^l + t_{rj}(y_{rj}^U - y_{rj}^l), r = 1, ..., s; j = 1, ..., n \text{ with } 0 \leq t_{rj} \leq 1, \]
\[ \omega_j = \omega_j^l + b_j(\omega_j^U - \omega_j^l), j = 1, ..., n \text{ with } 0 \leq b_j \leq 1. \]

Applying all the information and above transformation, the following models are obtained.

I. model – where the dual-role variable is not precisely known:

\[
\begin{align*}
\max f_o &= \sum_{r=1}^{s} \mu_r \left( y_{rjo}^l + t_{rjo}(y_{rjo}^U - y_{rjo}^l) \right) + \gamma \left( \omega_o^l + b_{jo}(\omega_o^U - \omega_o^l) \right) \\
&\quad - \beta \left( \omega_o^l + b_{jo}(\omega_o^U - \omega_o^l) \right), \\
\text{s.t.} &\quad \sum_{i=1}^{m} v_i (x_{ijo}^l + s_{ijo}(x_{ijo}^U - x_{ijo}^l)) = 1, \\
&\quad \sum_{r=1}^{s} \mu_r \left( y_{rj}^l + t_{rj}(y_{rj}^U - y_{rj}^l) \right) - \sum_{i=1}^{m} v_i (x_{ij}^l + s_{ij}(x_{ij}^U - x_{ij}^l)) \\
&\quad + \gamma \left( \omega_j^l + b_j(\omega_j^U - \omega_j^l) \right) - \beta \left( \omega_j^l + b_j(\omega_j^U - \omega_j^l) \right) \leq 0, \quad j = 1, ..., n \\
&\quad \mu_r \geq \varepsilon, 0 \leq t_{rj} \leq 1, \quad r = 1, ..., s \\
&\quad v_i \geq \varepsilon, 0 \leq s_{ij} \leq 1, \quad i = 1, ..., m \\
&\quad \gamma, \beta \geq \varepsilon, 0 \leq b_j \leq 1.
\end{align*}
\]

It can be seen that for inputs and outputs, there are new terms \( v_i s_{ij}, \mu_r t_{rj}, \gamma b_j \) and \( \beta b_j \), respectively. These new terms may be replaced by new variables \( q_{ij} = v_i s_{ij}, \quad p_{rj} = \mu_r t_{rj}, \quad c_j = \gamma b_j \quad \text{and} \quad d_j = \beta b_j \quad \text{which meet the needed conditions.} \) The model (13) can be rewritten as follows:

\[
\begin{align*}
\max f_o &= \sum_{r=1}^{s} \left( \mu_r y_{rjo}^l + p_{rjo}(y_{rjo}^U - y_{rjo}^l) \right) + \left( \gamma \omega_o^l + c_{jo}(\omega_o^U - \omega_o^l) \right) \\
&\quad - \left( \beta \omega_o^l + d_{jo}(\omega_o^U - \omega_o^l) \right) \\
\text{s.t.} &\quad \sum_{i=1}^{m} (v_i x_{ijo}^l + q_{ijo}(x_{ijo}^U - x_{ijo}^l)) = 1 \\
&\quad \sum_{r=1}^{s} \left( \mu_r y_{rj}^l + p_{rj}(y_{rj}^U - y_{rj}^l) \right) - \sum_{i=1}^{m} (v_i x_{ij}^l + q_{ij}(x_{ij}^U - x_{ij}^l)) \\
&\quad + \left( \gamma \omega_j^l + c_j(\omega_j^U - \omega_j^l) \right) - \left( \beta \omega_j^l + d_j(\omega_j^U - \omega_j^l) \right) \leq 0, \quad j = 1, ..., n \\
&\quad p_{rj} - \mu_r \leq 0, \quad r = 1, ..., s \\
&\quad q_{ijo} - v_i \leq 0, \quad i = 1, ..., m \\
&\quad c_j - \gamma \leq 0, \quad \text{for all } j \\
&\quad d_j - \beta \leq 0, \quad \text{for all } j \\
&\quad v_i, \mu_r, \gamma, \beta \geq \varepsilon, \quad \text{for all } i, r \\
&\quad q_{ij}, p_{rj}, c_j, d_j \geq 0, \quad \text{for all } i, r, j.
\end{align*}
\]
II. model - where the dual-role variable is precisely known:

\[
\max f_o = \sum_{r=1}^{s} (\mu_r y_{rjo}^L + p_{rjo} (y_{rjo}^U - y_{rjo}^L)) + (\gamma o - (\beta o))
\]

s.t.

\[
\sum_{i=1}^{m} (v_i x_{ijo}^L + q_{ijo} (x_{ijo}^U - x_{ijo}^L)) = 1
\]

\[
\sum_{r=1}^{s} (\mu_r y_{rj}^L + p_{rj} (y_{rj}^U - y_{rj}^L)) - \sum_{i=1}^{m} (v_i x_{ij}^L + q_{ij} (x_{ij}^U - x_{ij}^L)) + (\gamma j - (\beta j)) \leq 0, \quad j = 1, \ldots, n
\]

\[
p_{rj} - \mu_r \leq 0, \quad r = 1, \ldots, s
\]

\[
q_{ij} - v_i \leq 0, \quad i = 1, \ldots, m
\]

\[
v_i , \mu_r , \gamma, \beta \geq \varepsilon, \quad \text{for all } i, r, j.
\]

Variables in model (15) have the same meaning as in the I. model (14). Only just the dual-role variable is precisely known. The variable \( o \) is not in the interval.

Note, models (14) a (15) are models with the CRS assumption, but they can be easily extended to model with the VRS assumption, see the list of models in the end of the doctoral dissertation.

**Total factor productivity indices**

Generally, the productivity of one output and one input organization is defined as a ratio of output over input. The principle becomes more complex in the presence of multiple input/output case. Let there are \( n \) DMUs observed during \( T \) time periods with each using inputs \( x_j^t = (x_{1j}^t, x_{2j}^t, \ldots, x_{mj}^t) \) and producing outputs \( y_j^t = (y_{1j}^t, y_{2j}^t, \ldots, y_{sj}^t) \), where \( j = 1, 2, \ldots, n \) is a DMU index, \( t = 1, 2, \ldots, T \) denotes a respective time period, and \( m \) and \( s \) are the numbers of inputs and outputs, respectively. O'Donnell (2008) is formally defining the total factor productivity as followed:

\[
TFP_j^t = \frac{y_j^t}{x_j^t},
\]

(16)

where \( Y_j^t \equiv (y_j^t) \) is an aggregate output, \( X_j^t \equiv (x_j^t) \) is an aggregate input, and \( Y(\cdot) \) and \( X(\cdot) \) are non-negative, non-decreasing and linearly-homogenous aggregator functions, respectively. The associated index number that measures the TFP of the DMU \( j \) in period \( t \) with the TFP of DMU \( l \) in period \( r \) is following:

\[
TFP_{l,j}^{r,t} = \frac{TFP_l^t}{TFP_r^r} = \frac{y_j^t/x_j^t}{y_l^r/x_l^r} = \frac{y_j^t/y_l^r}{x_j^t/x_l^r} = \frac{y_j^t}{x_j^t},
\]

(17)

where \( Y_{l,j}^{r,t} \equiv y_j^t/y_l^r \) is an output quantity index and \( X_{l,j}^{r,t} \equiv x_j^t/x_l^r \) is an input quantity index. Thus, TFP growth can be expressed as a measure of output growth divided by a measure of input growth.
There are known different aggregator functions which give rise to different TFP index. If input and output prices are known, the aggregate quantities can be computed by employing Paasche, Laspeyres, Fisher, Tornqvist Indices. Otherwise, Malmquist, and Hicks-Moorsteen Indices relying on distance functions can be employed. However, O’Donnel (2011a) in this work have shown that all of these fail the transitivity test and thus cannot be used for multi-temporal and multilateral comparisons. Meanwhile, Lowe, Färe-Primont, and geometric Young Indices are suitable for such comparisons. Färe-Primont Index relies on distance functions and does not require price information. Indeed, it relies on non-linear weighting functions and normalised shadow prices (O’Donnell, 2011a). According to this information, the doctoral dissertation deals with the Malmquist index and the Hicks-Moorsteen index. and Färe-Primont index. All these indices do not need the input and output prices. The Malmquist index will be used to compare the results of this study with other studies. The Hicks-Moorsteen index will be used as new in this field and will be compared with the Malmquist index and each other to see if failing the transitivity assumption gives really different results for this case of study.

The class of non-negative, non-decreasing and linearly homogeneous aggregator functions for Malmquist\(r_l\) (18), Malmquist\(t_j\) (19), Hicks-Moorsteen (20), respectively, are following:

\[Y(y) = D_O(x_l^i, y, r)\] and \[X(x) = D_I(x, y_l^j, t),\]
\[Y(y) = D_O(x_j^i, y, t)\] and \[X(x) = D_I(x, y_j^j, t),\]
\[Y(y) = [D_O(x_l^i, y, r), D_O(x_j^j, y, t)]^{1/2}\] and
\[X(x) = [D_I(x, y_l^i, r), D_I(x, y_j^j, t)]^{1/2},\]

and where \(x_0\) and \(y_0\) are vectors of quantities, \(t_0\) denotes a representative time period and \(D_O(\cdot)\) and \(D_I(\cdot)\) are Shephard (1972) output and input distance functions. The aggregator functions of inputs and outputs (18) to (20) are so-named because they are substituted into (16) and (17) the give rise to the following TFP indices:

\[TFP_{l,j}^{r,t} = \frac{D_O(x_l^i, y_l^j, r)}{D_O(x_l^i, y_l^j, r)} \frac{D_I(x_l^i, y_l^j, r)}{D_I(x_l^i, y_l^j, r)},\]
\[TFP_{l,j}^{r,t} = \frac{D_O(x_j^j, y_j^j, t)}{D_O(x_j^j, y_j^j, t)} \frac{D_I(x_j^j, y_j^j, t)}{D_I(x_j^j, y_j^j, t)},\]
\[TFP_{l,j}^{r,t} = \left(\frac{D_O(x_l^i, y_l^j, r)}{D_O(x_l^i, y_l^j, r)} \frac{D_I(x_l^i, y_l^j, r)}{D_I(x_l^i, y_l^j, r)} \frac{D_O(x_j^j, y_j^j, t)}{D_O(x_j^j, y_j^j, t)} \frac{D_I(x_j^j, y_j^j, t)}{D_I(x_j^j, y_j^j, t)}\right)^{1/2},\]

where (21) is Malmquist\(r_l\), (22) is Malmquist\(r_j\), (23) is Hicks-Moorsteen.

It refers to the indices (21) and (22) as Malmquist\(r_l\) and Malmquist\(r_j\) indices because the component output quantity and input quantity indices are the organization-specific Malmquist indices defined by Caves et al. (1982). Färe and other authors (e.g., Färe et al. (1997), Ray and Desli (1997)) made further development of this index. The index defined by (23) was first proposed by Bjurek (1996) but is commonly known as a Hicks-Moorsteen index because it is the geometric average of two indices that Diewert (1992) attributed to Hicks (1961) and Moorsteen (1961).

Note, all the models had been calculated by the author in software General Algebraic Modeling System (GAMS).
6 Summary of results and conclusion

The financial sector is an important part of today’s economic system. The financial sector has an extreme impact on both macroeconomics and microeconomics. Banking and other financial institutions play a significant role. For example, they affect the allocation of financial resources. Banks and financial institutions have to be efficient if they want to properly allocate the economic resources. Nowadays, the global interconnection of all participants in the financial sector brings both advantages (easier mobility capital, more efficient allocation of resource, lower prices for consumers and so on) and disadvantages. If the economy of some country or region is not healthy, the impact of the unhealthy and inefficient financial sector may have a great effect on each individual participant and also on the entire economic system.

That is why monitoring, analysing and better understanding of the situation in the banking and financial sector should lead to minimizing the problems. More precisely, efficiency measurements of the banking and financial sector determine the performance of all evaluated units (banks and financial institutions) and it helps them to detect different possibilities to improve. These measurements provide valuable knowledge to the market regulators and give them material for their decisions. The responsible institutions are able to set better rules for management and supervision in the financial sector if they know the real state of the efficiency score compared to others.

This doctoral dissertation focuses on the creation of new DEA model/models which will serve to better monitor, analyse and understand to the situation in the financial sector of the VG countries. In other words, the main goal of the doctoral dissertation is to create or modify the basic DEA models based on the needs of the financial environment, and apply these new models to evaluate the financial sector in the Visegrad Group countries during the time period 2005-2015. The theoretical part of this doctoral dissertation is done in Chapter 3. There are introduced a two new models - model M6 - Model with Risk Variable I and model M7 - Model with Risk Variable II. The practical part is solved in Chapter 4.

Four scientific questions are defined to achieve the main aim. These scientific questions are also solved in this doctoral dissertation. All these questions are solved mathematically in Chapter 4 and applied in Chapter 5. However, before their mathematical solution, it is necessary to solve the model variables and the assumption of return to scale. For the selection of variables, a classic intermediation approach and a model that is extended by non-traditional activities is used. The second mentioned model with non-traditional activities is found to be more relevant to the issue. The assumption of return to scale is also drawn and there is seen a first signs of the VRS confirmation, but still both versions are tested in the following first scientific question. The first scientific question deals with the general problem of the financial institutions - missing data and that means to deal with the DEA models for missing variables. The problem of missing parameters is solved by DEA models with data in the interval. These intervals are obtained by several estimates - regressions, averages, etc. Based on the close analysis of all different models with missing data and a different assumption for the return to scale (CRS and VRS) or a different set of input and output variables, there is found one, the best version of the model - the model with the VRS assumption. This model is used and modified again to solve the second scientific question. More specifically, how to deal with the idea that some input or output variables may be in both positions (as input and output variables in one time - in this case the deposits). The special DEA model with the dual-role variable has been managed. This model has shown that the deposits should really be on the input side. However, this improved DEA model
with the dual-role variable has been used in the third scientific question for another modification. This question deals with the problem of the risk variable - a variable which may have a different position, but the goal is to put this special variable into the model to provide more information about the environment of the financial institutions and risks which may occur in the region. Based on these three scientific question, two new models are defined in subsection 5.5. One of this models is used farther for the last scientific question. Which of the two indices (the classical total productive index (Malmquist) or the alternative total productive index (Hicks)) gives more information and has better tangible values. The last scientific question has not been completely solved. Based on previous research, it was assumed that the alternative index should be more advantageous, but with the use of the new model, there was no such calculation. In the end, the results of a new calculation were the slightly similar - the alternative index provides the lower score then the classical, but the trend has been same. However, the new models have met expectations and they have expanded opportunities for better understanding the situation of the analysis in the VG countries and their financial sector.

In the future, it would be appropriate to extend the database of financial institutions in terms of time and quantity. It would also be worthwhile exploring other options for dual-role variables. The possibility of using deposits as a dual-role variable (as variable of input and output position) is contemplated in the present case based on the percentage distribution, which should be related to the size of the financial institutions. Extending other risk factors would also be appropriate, for example as some fuzzy variables or VaR. Also, the fourth scientific question has to be analysed closely.
7 List of references


8 List of author’s publications and research


The subject of this doctoral dissertation is the issue of measuring the efficiency of financial institutions in the Visegrad Group (VG) countries. The main aim of the doctoral dissertation is to create and modify the basic CCR and BCC Data Envelopment Analysis (DEA) models and apply these new models to the financial sector of the VG countries in the reference time period 2005-2015. The fulfilment of the goal of the doctoral dissertation is closely related to the solution of four research questions which focus and solve problems or shortcomings of models and the environment of financial institutions in the VG countries. These problems include the lack of data (e.g., the problem of missing values in a number of input and output variables), the problem of clearly defining variables as input or output variables or the use of risk variables that help to better illustrate the financial institutions' environment and offset differences between VG countries. Finally, the answer to the question is sought, whether an alternative productive index (Hicks index) is more suitable than a traditional productive index (Malmquist index) for a given specific area.

The creation and modification of models took place in analogously with defined research questions as well as their mathematical application. More precisely, an analysis of basic DEA models with different types of input and output variables had been performed first (both in terms of defining inputs and outputs, and the possibility of replacing the missing variables). On the basis of the analyses, the best model was determined, and it was further tested and modified by using the dual variable and the risk variable. All models were applied, verified and the best model was subsequently used as the basis for the Hicks and Malmquist index. These indices have been compared and found to give a more suitable description of given area.

The result of the doctoral dissertation is two new models, which should correspond to the conditions and needs of financial institutions in the VG countries. They should measure the efficiency better. This means providing a more accurate and deeper analysis in the area. It can help to these institutions to find their weaknesses and improve them in future.

Key words:

Models of Data Envelopment Analysis (DEA models), efficiency, Visegrad Group (VG) countries, missing values, dual-role variables, risk variables, total factor productivity index, Hicks index, Malmquist index.
10 Shrnutí

Předmětem řešení doktorské dizertační práce je problematika měření efektivnosti finančních institucí v zemích Visegrádské skupiny. Hlavním cílem doktorské dizertační práce je vytvořit a modifikovat základní CCR a BCC modely analýzy obalu dat (DEA modely) a aplikovat tyto nové modely na výše zmíněný finanční sektor zemí Visegrádské skupiny v referenčním období 2005-2015. Naplnění cíle doktorské dizertační práce úzce souvisí s řešením čtyř výzkumných otázek, které se zaměřují a řeší problematiku či nedostatky modelů a prostředí finančních institucí v oblasti Visegrádské skupiny. Mezi zmíněné problémy patří nedostatek dat (např. problém chybějících hodnot v množině vstupů a výstupů), problém jasného definování proměnných jako vstupní či výstupní proměnné nebo možnost použití rizikových proměnných, které pomáhají lépe dokreslovat prostředí finančních institucí a vyrovnávat rozdíly mezi zeměmi Visegrádské skupiny. V neposlední řadě je hledána odpověď na otázku, zda alternativní produktivní index (Hicksův index) není vhodnější pro danou specifickou oblast, než klasický index produktivity (Malmquistův index).

Vytvoření a modifikování modelů probíhalo v analogii s definovanými výzkumnými otázkami, stejně jako jejich matematická aplikace. Přesněji, nejprve byla provedena analýza základních DEA modelů s různými druhy proměnných, jak z hlediska definování vstupů a výstupů, tak možnosti nahrazení chybějících proměnných. Na základě analýzy byl stanoven nejlepší model, a ten byl dále testován a modifikován za pomoci duální proměnné a rizikové proměnné. Všechny modely byly aplikovány, ověřovány a nejlepší model byl následně použit jako základ pro Hicksův a Malmquistův index. Tyto indexy byly porovnány a bylo zjištěno, který pro danou oblast dává lepší popis situace.

Výsledkem disertační práce jsou dva nové modely, které by měly odpovídat podmínkám a potřebám finančních institucí v prostředí Visegrádské skupiny a dokázat lépe měřit efektivnost v této oblasti, což může pomoci těmto institucím najít své slabé stránky a dále se zlepšovat.

Klíčová slova:

Modely analýzy obalu dat (DEA modely), efektivita, Visegrádská skupina (VG), chybějící hodnoty, proměnná dvoji role, riziková proměnná, index celkové produktivity, Hicksův index, Malmquistův index.