Andrea Furková
Department of Operations
Research and Econometrics
Faculty of Economic Informatics,
University of Economics in Bratislava
Dolnozemská cesta 1/b
852 35 Bratislava

APPLICATION OF PARAMETRIC AND NON-PARAMETRIC BENCHMARKING METHODS IN COST EFFICIENCY ANALYSIS OF THE ELECTRICITY DISTRIBUTION SECTOR

Summary

This paper explores the application of parametric and non-parametric benchmarking methods in measuring cost efficiency of Slovak and Czech electricity distribution companies. We compare the relative cost efficiency of Slovak and Czech distribution companies using two benchmarking methods: the non-parametric Data Envelopment Analysis (DEA) and the Stochastic Frontier Analysis (SFA) as the parametric approach. The first part of analysis was based on DEA models. Traditional cross-section CCR and BCC model were modified to cost efficiency estimation. In further analysis we focus on two versions of stochastic frontier cost function using panel data: MLE model and GLS model. These models have been applied to an unbalanced panel of 11 (Slovakia 3 and Czech Republic 8) regional electricity distribution utilities over a period from 2000 to 2004. The differences in estimated scores, parameters and ranking of utilities were analyzed. We observed significant differences between parametric methods and DEA approach.

Key words

Data Envelopment Analysis, Stochastic Frontier Analysis, Cost Efficiency, Panel Data Models, Electricity Distribution Sector

ACM classification

J.4 SOCIAL AND BEHAVIORAL SCIENCES, *Economics,* G.1 NUMERICAL ANALYSIS, *Linear programming*

JEL classification

C6 – Mathematical Methods and Programming, C61 – Optimization Techniques; Programming Models; Dynamic Analysis

1. INTRODUCTION

Electricity sector reforms are transforming the structure and operating environment of the electricity industries across many European countries. The main aims of these reforms are to introduce market-oriented principles in generation, transmission and distribution of electricity and to increase the efficiency of natural monopoly services by introducing the „right“ regulation schemes.

Slovakia and all east European countries having joined the EU, have to apply recent EU directives in particular the Electricity Directive 2003/54/EC as the key European legislation establishing the internal market of electricity. The main objective of this directive is to improve efficiency and competitiveness of power sector. As the power sector is specific sector and the nature of network industries utilities does not allow introduction of full competition, it is still necessary to regulate this sector.

In Slovakia a role of independent regulator from 2001 performs Agency for network industries regulation (ÚRSO). From 2003 this agency performs price cap regulation by publishing decrees and resolutions.

Regulated prices must involve reasonable profit and economically substantial costs. Market liberalization of Slovak power sector is delayed: in 2005 only entrepreneurs became eligible customers and in 2007 power generation and electricity supply should become competitive activities.

Two main regulation schemes of distribution utilities are Rate of Return Regulation and Price Cap Regulation. Nowadays many regulators have adopted incentive-based regulation models instead of traditional rate of return regulation. In traditional rate of return regulation systems companies recover their costs with a risk-free rate of return and therefore have little incentive to minimize costs. On the other hand, price cap regulation is incentive - based scheme and these kinds of regulation schemes are designed to provide incentive for efficiency improvement. Price cap regulation is based on the (RPI – X) (inflation factor-efficiency factor) formula.

Regulation of Slovak and Czech distribution utilities is currently following incentive price cap regulation and the main issue of this regulation scheme is how to set efficiency factor X? A widely applied approach in many European countries is benchmarking, that is measuring a company's efficiency compared with a reference performance (so-called efficient frontier). Inefficiency can results from technological deficiencies or non-optimal allocation of resources into production. Both technical and allocative inefficiencies are included in cost-inefficiency. Many studies in this field used the physical quantities of inputs measures for efficiency improvements in terms of reductions in physical units. Nevertheless, this could be in conflict with the main goal of regulator, which is cost reduction of regulated utilities. Therefore we decided to orient this study to less developed and applicated problem of cost efficiency prediction. As an analogous efficiency analysis of Slovak and Czech electricity distribution utilities was not done yet and the methodology of selected parametric (SFA) and non-parametric (DEA) methods for cost efficiency estimation is not in Slovak literature sufficiently elaborated we briefly present in next two chapters relevant methodology.

Generally, there are two families of methods used in benchmarking praxis:

- **Non-parametric methods**, like Data Envelopment Analysis (DEA), originate from Operations Research and use linear programming to calculate an efficient deterministic frontier against which companies are compared. Detail presentation of DEA models can be found in [Cooper, Seiford and Tone (2000)] and [Coelli, Rao Prasada and Battese (1998)].
- **Parametric methods**, like Stochastic Frontier Analysis (SFA), use econometric theory to estimate pre-specified functional form and inefficiency is modelled as an additional stochastic term. Detail presentation of SFA models can be found in [Kumbhakar and Lovell (2003)] and [Coelli, Rao Prasada and Battese (1998)].

2. DATA ENVELOPMENT ANALYSIS – NONPARAMETRIC BENCHMARKING TECHNIQUE

DEA method is the most commonly used approach in practice. After setting strategic behaviour to cost minimization traditional input oriented CCR (with constant returns to scale – CRS) and BCC (with variable returns to scale – VRS) models for a sample of N companies with a *k*-input-*m*-output production function can be modified to cost efficiency estimation problem:

$$\min_{\boldsymbol{\lambda},\mathbf{x}_i^*} \mathbf{w}_i^T \mathbf{x}_i^* \qquad (1)$$

$$-\mathbf{y}_i + \mathbf{Y}\boldsymbol{\lambda} \geq \mathbf{0}$$

$$\mathbf{x}_i^* - \mathbf{X}\boldsymbol{\lambda} \geq \mathbf{0}$$

$$\boldsymbol{\lambda} \geq \mathbf{0}$$

where

w_i and x_i are *k*x1 vectors respectively representing input prices and quantities for firm *i*,

y_i is an *m*x1 vector representing the given output,

X and Y are respectively *k*xN input and *m*xN output matrices,

λ is an Nx1 vector of non-negative constants to be estimated.

In BCC model the VRS property is satisfied through the convexity constraint $\mathbf{e}^T \boldsymbol{\lambda} = 1$.

The solution of minimization problem in (1) gives the minimum feasible costs for each company namely, $\mathbf{w}_i^T \mathbf{x}_i^*$, where $\mathbf{x}_i^*$ is the optimal input vector for firm *i*. The cost-efficiency of each company is then estimated as its distance to the envelope, i.e. as ratio of the minimum feasible costs to actual costs (for more details see [Coelli, Rao Prasada and Battese (1998)]:

$$CE = \mathbf{w}_i^T \mathbf{x}_i^* \Big/ \mathbf{w}_i^T \mathbf{x}_i \qquad (2)$$

3. STOCHASTIC FRONTIER ANALYSIS – PARAMETRIC BENCHMARKING TECHNIQUE

Further analysis was focused on the parametric approach – SFA. In this approach is used econometric theory to estimate pre-specified functional form and inefficiency is modelled as an additional stochastic term. Frontier cost function identifies the minimum costs at a given output level, input prices and existing production technology.

The stochastic frontier cost function (single Cobb-Douglas form) for panel data is formulated as:

$$\ln C_{it} = \beta_0 + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit} + v_{it} + u_i$$

$$i = 1,2,\ldots,N \text{ and } t = 1,2,\ldots,T \qquad (3)$$

where

C_{it} are observed total costs of the i-th firm in year t,

y_{it} is a vector of outputs of the i-th firm in year t,

w_{it} is an input price vector of the i-th firm in year t,

u_i are non negative time-invariant random variables assumed to be half normal distributed ($u_i \sim \text{iid}N^+(0,\sigma_u^2)$),

v_{it} are random variables which are assumed to be normally distributed ($vit \sim iidN(0,\sigma_v^2)$).

In this specification the error term is composed of two uncorrelated parts. The first part u_i is capturing the effect of inefficiency (including both allocative and technical inefficiencies) and the second part v_{it} is reflecting effect of statistical noise. This random effect model can be estimated using Maximum Likelihood Estimation (MLE) method and Battese and Coelli point estimator (for more details see [Kumbhakar and Lovell (2003)]) can be used for estimation of inefficiency scores.

The main advantage of the stochastic cost frontier approach is the separation of the inefficiency effect from the statistical noise. However, this model is subject to the potential criticism of having an arbitrary assumption about the distribution of the random terms.

Another random effect model was proposed by Schmidt and Sickles to overcome the problem of specifying a particular distribution for the inefficiency by rewriting equation (3):

$$\ln C_{it} = \beta_0 + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit} + v_{it} + \alpha_i$$

$$i = 1,2,\ldots,N \text{ and } t = 1,2,\ldots,T \qquad (4)$$

where

$\alpha_i = \alpha + u_i$ and α is an intercept. Now, the Generalized Least Squares (GLS) method can be used. The estimate of inefficiency component is defined as the distance from the firm specific intercept to the minimal intercept in the sample:

$$u_i = \alpha_i - \min_i(\alpha_i) \qquad (5)$$

There are no required assumptions about distribution for inefficiency in this model. The remaining restrictive assumption is that two random components must not be correlated with each of the explanatory variables.

3. MODEL SPECIFICATION AND DATA

The above-mentioned models have been applied to an unbalanced panel of 11 (Slovakia 3 and Czech Republic 8) regional electricity distribution utilities over a period from 2000 to 2004. The sample includes 54 observations (missing all data for Jihočeská energetika for year 2000). All data are based on information from the annual reports of distribution companies. Selected measures for Slovak and Czech distribution utilities are given in table 1.

The first part of analysis was based on DEA models, followed by analysis SFA models. We applied cross-section CCR and BCC model (year 2004) and two versions of SFA panel data models discussed in chapter two and three in order to achieve scores of cost efficiencies and analyze the differencies across models. A similar benchmarking analysis of Swiss electricity distribution utilities was done by authors Farsi and Filippini [Farsi and Filippini (2005)], but there were only cross data used in it.

3.1 DEA MODELS

In DEA method there is no need to specify any functional form. We applied traditional cross-section (year 2004) CCR with constant

Table 1 **Descriptive statistics (54 observations)**

	Mean	Median	Std. Dev.	Minimum	Maximum
Total annual costs (*C*) in mil. SKK	14876	14446	3824	7819	21308
Annual output (*Y*) in GWh	6051	6289	1481	3368	8840
Average capital price (P_K) in thous. SKK per MVA of installed capacity	1740	1469	815	887	4501
Average annual labour price (P_L) per employee in SKK	342516	342992	79151	204423	628603
Average price of input power (P_P) in SKK/MWh	1547	1535	200	1077	2003
Number of customers (*CU*)	711702	666006	182709	401183	1018558
Service area (*AS*) in km sq.	11619	11242	4541	500	17978
Customer density (*CUD*)	179	60	383	35	1394

Source: Annual reports 2000 – 2004

return to scale and BCC model with variable returns to scale. After setting strategic behaviour (cost minimization) CCR and BCC model were modified to cost efficiency estimation. As input variables in both DEA models have been chosen:

- labour (*L*) – was defined as the average annual number of utility's employees
- capital (*K*) - was defined as the total installed capacity of the utility's transformers in MVA
- purchased energy (*P*) – was defined as the total purchased energy from the generator in MWh
- labour price (P_L) – is defined as the average annual salary of the utility's employees
- capital price (P_K) – is measured as the ratio of capital expenses to the total installed capacity of the utility's transformers in MVA
- purchased energy price (P_P) – is defined as average price of purchased energy from generator.

As output variable have been chosen:

- total output (*Y*) – was measured as the total number of delivered electricity in MWh.

The minimization problem given in (1) can be solved by linear programming method. This method finds a piece-wise linear isoquant in the input space, which corresponds to the minimum costs of producing the given output at any given point. The solution gives the minimum feasible costs for each company namely, $w^T x_i^*$, where x_i^* is the optimal input vector for firm *i*. The cost efficiency of each company is then estimated as its distance to the envelope. For the estimation of the efficient envelope in our case is necessary to solve 11 linear programming problems given in (1). For example, for the first distribution utility (Pražská energetika) according to CCR model we formulated following minimization problem:

$$\min_{\lambda, x_i} 1228x_1^* + 428397x_2^* + 1606x_3^*$$

$$-5260 + \begin{pmatrix} 5260\lambda_1 + 8019\lambda_2 + 3739\lambda_3 + 8296\lambda_4 + \\ +6213\lambda_5 + 4402\lambda_6 + 6196\lambda_7 + \\ +6438\lambda_8 + 7299\lambda_9 + 4815\lambda_{10} + 6315\lambda_{11} \end{pmatrix} \geq 0$$

$$x_1^* - \begin{pmatrix} 2476\lambda_1 + 3368\lambda_2 + 1570\lambda_3 + 3642\lambda_4 + \\ +2486\lambda_5 + 1849\lambda_6 + 2663\lambda_7 + 2060\lambda_8 + \\ +2336\lambda_9 + 1541\lambda_{10} + 1992\lambda_{11} \end{pmatrix} \geq 0$$

$$x_2^* - \begin{pmatrix} 1238\lambda_1 + 1610\lambda_2 + 1120\lambda_3 + 1612\lambda_4 + \\ +1466\lambda_5 + 1529\lambda_6 + 1430\lambda_7 + 1354\lambda_8 + \\ +1426\lambda_9 + 1845\lambda_{10} + 2085\lambda_{11} \end{pmatrix} \geq 0$$

$$x_3^* - \begin{pmatrix} 5813\lambda_1 + 9210\lambda_2 + 4025\lambda_3 + 8998\lambda_4 + \\ +7038\lambda_5 + 4651\lambda_6 + 6719\lambda_7 + 6872\lambda_8 + \\ +7816\lambda_9 + 5225\lambda_{10} + 6874\lambda_{11} \end{pmatrix} \geq 0$$

$$\lambda_{1,2,3,4,5,6,7,8,9,10,11} \geq 0 \qquad (6)$$

The solution gives the optimal input values x^* (x_1^* - capital, x_2^* - labour, x_3^* - purchased energy - are for each distribution utility listed in table 2), which are necessary for cost efficiency estimation of each utility. The cost efficiency is defined as ratio of the minimum feasible costs to actual costs, for example for the first distribution utility (Pražská energetika) was calculated as follows:

$$CE = \frac{1228 * 2309{,}18 + 428397 * 1022{,}07 + 1606 * 5705{,}1}{1228 * 2476 + 428397 * 1238 + 1606 * 5813} = \\ =0{,}829$$

The estimation of efficient frontier in BCC model is analogous as in CCR model[1]. The VRS property in BCC model is satisfied through the convexity constraint $\mathbf{e}^T \lambda = 1$.

Table 2: **Actual and optimal levels of inputs**

Company	Y	Actual levels of inputs			Optimal levels of inputs according to CCR model			Optimal levels of inputs according to BCC model		
		K	L	P	K	L	P	K	L	P
Pražská energetika	5260	2476	1238	5813	2309,18	1022,07	5705,10	2476,00	1238,00	5813,00
Jihomoravská energetika	8019	3368	1610	9210	3520,40	1558,18	8697,56	3279,15	1560,32	8669,60
Jihočeská energetika	3739	1570	1120	4025	1641,45	726,53	4055,39	1570,00	1120,00	4025,00
Severomoravská energetika	8296	3642	1612	8998	3642,00	1612,00	8998,00	3642,00	1612,00	8998,00
Stredočeská energetika	6213	2486	1466	7038	2727,55	1207,25	6738,74	2410,57	1325,87	6749,17
Západočeská energetika	4402	1849	1529	4651	1408,83	860,02	4713,80	1964,92	1171,44	4804,39
Východočeská energetika	6196	2663	1430	6719	2720,09	1203,95	6720,30	2411,73	1324,30	6732,47
Severočeská energetika	6438	2060	1354	6872	2826,33	1250,97	6982,78	2395,12	1346,61	6970,20
Západoslovenská energetika	7299	2336	1426	7816	3204,31	1418,27	7916,64	2336,00	1426,00	7816,00
Východoslovenská energetika	4815	1541	1845	5225	1541,01	940,70	5156.05	2210,93	1203,48	5289,88
Stredoslovenská energetika	6315	1992	2085	6874	2772,33	1227,07	6849,37	2403,56	1335,27	6849,37
Average	6090	2362	1520	6658	2574	1184	6738	2464	1333	6611

Source: Self calculations

3.2 SFA MODELS

To illustrate differences across models, we focus on two versions of SFA models, namely ML model and GLS model. In both models we used panel data set for 3 Slovak and 8 Czech electricity distribution utilities over the 2000 - 2004 period and we adopted traditional Cobb-Douglas functional form. A triple-input and single-output Cobb-Douglas cost function has been considered. The output is measured as the total number of delivered electricity in MWh, and the three input factors are set as capital, labour and the input power purchased from generator. Capital price is measured as the ratio of capital expenses to the total installed capacity of the utility's transformers in MVA. The capital costs are approximated by the residual costs, i.e. total costs minus labour and purchased power costs. Labour price is defined as the average annual salary of the firm's employees. As electricity distribution utilities operate in networks with different shapes, which directly affect the costs, the cost function should take into account differences in network characteristics. For this reason we included in the cost function such factor that is unrelated to cost-efficiency but affects the costs. In addition to input prices and output, factor *CUD* - customer density (defined as number of customers per km sq. service area) is included. The condition of linear homogeneity in input prices was imposed by dividing money values by the price of the input power. The Cobb-Douglas specification of the cost function can be then formulated as follows:

where C represents total costs, Y is the output, P_K, P_L, P_P are the prices of capital, labour and input power respectively, CUD is customer density and v_{it} and u_i are random variables described in equations (3) and (4). Both models (MLE model, GLS model) are based on specification given in (7). The differences are in the specification of the residuals. This term is composed of two components, one of which (u_i) being time-invariant (firm-specific) and the other (v_{it}) varying across observations. In the case of MLE model, the cost function in (7) has been estimated by Maximum Likelihood Estimation method[2] and in the case of GLS model we used Generalized Least Squares method[3].

$$\ln\left(\frac{C}{P_P}\right)_{it} = \beta_0 + \beta_Y \ln Y_{it} + \beta_K \ln\left(\frac{P_K}{P_P}\right)_{it} + \beta_L \ln\left(\frac{P_L}{P_P}\right)_{it} + \beta_{CUD} \ln CUD_{it} + v_{it} + u_i$$
$$i = 1,\ldots,N \quad t = 1,\ldots,T \qquad (7)$$

2 We used Frontier 4.1

3 We used Eviews5

4. ESTIMATION RESULTS

The estimated parameters of the cost frontier are listed in table 3. This table shows that almost all the coefficients are significant and the coefficients are not significantly different from one model to another, suggesting the results for the parameters do not depend on distributional assumptions of the error and inefficiency term and method of estimation.

Table 3 **Cost frontier parameters**

	MLE – model		GLS – model	
	Coeff.	Std. Error	Coeff.	Std. Error
Constant	– 5,7196*	0,3038	-5,6117*	0,3583
ln*Y*	0,8673*	0,0337	0,8565*	0,0433
lnP_K/P_P	0,2645*	0,0134	0,2667*	0,0115
lnP_L/P_P	0,0684*	0,0217	0,0774*	0,0215
ln*CUD*	0,0010	0,0088	-0,0022	0,0139
s	0,0045	0,0020		
g	0,8475	0,0799		
R-squared			0,9603	

* significant at p=0,05

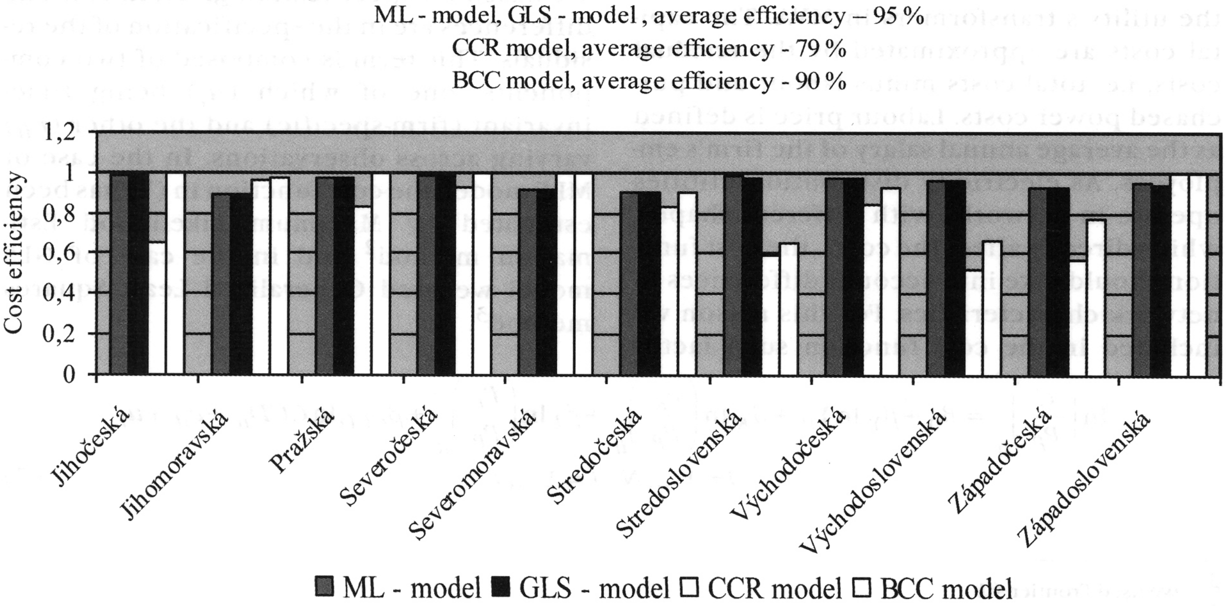
Table 4 provides efficiency estimates and ordering of companies according to MLE model, GLS model, CCR model and BCC model (the estimated efficiency scores are given in brackets). The scores can move between 0 and 1, where the highest value implies a perfectly efficient company and the difference from 1 approximates the percentage of the total

Table 4 **Efficiency ranking for the companies and efficiency scores**

Company	MLE - model	GLS - model	CCR - model	BCC - model
Severočeská energetika	1 (0,9915)	1 (1,0000)	4(0,927)	5(0,996)
Jihočeská energetika	2 (0,9871)	1 (1,0000)	8(0,655)	1(1,000)
Stredoslovenská energetika	3 (0,9815)	3 (0,9783)	9(0,600)	11(0,650)
Pražská energetika	4 (0,9788)	5 (0,9694)	6(0,829)	1(1,000)
Východočeská energetika	5 (0,9756)	4 (0,9726)	5(0,846)	7(0,927)
Východoslovenská energetika	6 (0,9646)	6 (0,9642)	11(0,519)	10(0,663)
Západoslovenská energetika	7 (0,9531)	7 (0,9433)	2(0,999)	1(1,000)
Západočeská energetika	8 (0,9284)	8 (0,9320)	10(0,569)	9(0,772)
Severomoravská energetika	9 (0,9174)	9 (0,9099)	1(1,000)	1(1,000)
Stredočeská energetika	10 (0,9041)	10 (0,8983)	7(0,827)	8(0,906)
Jihomoravská energetika	11 (0,8945)	11 (0,8866)	3(0,968)	6(0,969)
Average	**(0,9524)**	**(0,9504)**	**(0,794)**	**(0,898)**

Source: Self calculations

Figure 1 **Efficiency ranking for the utilities and efficiency scores, SFA (ML – model, GLS – model) and DEA (CCR – model, BCC – model)**



Source: Self calculations

costs that the company can potentially save. As the results suggest, the studied companies are on average about 95 % efficient according to SFA models. The individual efficiency estimates are stable across models. The studied companies are on average about 79 % efficient according to CCR model and 90 % efficient according to BCC model.

The individual estimated cost efficiency scores according to all models also provides next figure.

5. CONCLUSIONS

We have found out that efficiency scores, the estimated parameters of cost function and ranks are robust on estimation procedure in SFA models (GLS – model, ML – model). In this study it was not possible to estimate cross – section SFA model (only 11 observations – due to the small number of utilities in the sector) and therefore we had to compare SFA panel data models with DEA cross section data models. This could be possible source of differences in the results (efficiency scores and ranks) between parametric method (SFA) and DEA approach and therefore in our opinion a benchmarking analysis can be used by regulator as an auxiliary instrument to establish a larger informational basis for more effective price cap regulation, but the results should be used with caution since the results can be influenced by the method and the model specification.

REFERENCES

1. Annual reports companies 2000 – 2004
2. Coelli, T.: A Guide to Frontier Version 4.1, Centre for Energy Policy and Economics, Swiss Federal Institutes of Technology, Zurich, 1996
3. Coelli, T. J.: A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program, Working Paper, Centre for Efficiency and Productivity Analysis, University of New England, Australia, 1996
4. Coelli, T. J., Rao Prasada, D., Battese, G. (1998): An Introduction to Efficiency and Productivity Analysis, Kluwer Academic Publishers.
5. Farsi, M., Filippini, M.: A Benchmarking Analysis of Electricity Distribution Utilities in Switzerland, Centre for Energy Policy and Economics, Swiss Federal Institutes of Technology, Zurich, 2005
6. Greene, W., Farsi, M., Filippini, M: Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector, Centre for Energy Policy and Economics, Swiss Federal Institutes of Technology, Zurich, 2005
7. Kumbhakar, Subal C., Lovell, C. A. Knox: Stochastic Frontier Analysis, Cambridge University Press, Cambridge, 2003
8. Cooper, W. W., Seiford, L. M., Tone, K.: Data Envelopment Analysis, Kluwer Academic Publishers, 2000

Biography:

Ing. Andrea Furková, PhD. is a teacher of the Department of Operations Research and Econometrics at the Faculty of Economic Informatics since 2000. Her research activities are focused on parametric and non-parametric benchmarking methods in cost efficiency estimation.