Cobb-Douglas or Translog Production Function in Efficiency Analysis?

Marie Šimpachová Pechrová[1], Ondřej Šimpach[1]

Abstract. The aim of this paper is to assess two types of specification of production function when estimating the technical efficiency with Stochastic Frontier Analysis – Cobb-Douglas (CD) and translog (TL) production functions and their influence on the results. CD production function is easy to be estimated. It is a special case in a more general class of production functions with constant elasticity of substitution, where is the mostly used a TL function.

We compared both specifications on the unbalanced panel data of agricultural holdings in the Czech Republic. There were 517 farms and 1708 observations for years 2013 to 2016. Accounting data were taken from Albertina database purchased from Bisnode, s. r. o., company and data about acreage were taken from Land Parcel Identification System. Production was approximated by sales, production factors were: consumption of material and energy, fixed assets, number of employees and acreage of agricultural land. Technical efficiencies were calculated by JLMS (Jondrow, Lovell, Materov, Schmidt, 1982 [11]) method and their differences tested by Wilcoxon signed-rank test and correlated by Spearman's rank coefficient.

The technical efficiency in CD specification was slightly higher (85.69%) than in TL specification (85.12%), but there were found statistically significant differences. Spearman's $\rho = 0.9597$ was close to 1 pointing on high positive dependence. Using likelihood ratio test, the CD specification was rejected in favor of more general specification of TL function.

We can conclude that TL function is more appropriate to be used as a production function in SFA, but the effect on the technical efficiency is mild and almost negligible (despite that the test revealed statistically significant differences, the dependence was very high.)

Keywords: agricultural holdings, Cobb-Douglas, production function, stochastic frontier analysis, translog production function

JEL Classification: C23, C52
AMS Classification: 90C15

1 Introduction

The concept of technical efficiency was firstly introduced by Farrell. The idea is to compare so-called decision- making units according to their ability to transform the inputs to outputs. Those units (typically firms) that are able to produce the given output with the lowest possible combination of inputs (input-oriented models) or are able to attain the highest possible output with given inputs (output-oriented models) are laying on the frontier and are 100% efficient. Other firms are less technically efficient and shall improve their performance. There are two approaches to efficiency measurement: non-parametric and parametric. The well-known are two representants of each group – Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA).

Data envelopment analysis was first proposed by Charnes et al. [6] "It uses linear programming techniques to build a non-parametric efficiency frontier of the data sample." [22]. It can include more inputs and outputs, but its usage on panel data is limited. As we have unbalance panel of agricultural holdings, we rather use parametric method, particularly SFA.

The SFA was simultaneously developed by Meeusen and Van den Broek [15] and Aigner et al. [1]. It estimates a parametric frontier of the best possible practices given a standard cost or profit function. [22]. Its advantage is that it can be used on panel data, can include determinants of technical inefficiency in the

[1] Institute of Agricultural Economics and Information, Mánesova 75, Prague 2, simpachova.marie@uzei.cz, simpach.ondrej@uzei.cz

model and assess them in one step, and takes into account statistical noise, so the differences among firms are given not only due to inefficiency, but also because of other stochastic factors. On the other hand, it is necessary to correctly specify the production technology and the distribution of the inefficiency term which is not always easy. Inefficiency term can have exponential, gamma, half-normal or truncated normal distribution. "Notice that the half-normal is a special case of the truncated normal." [13]. Since the SFA measures the efficiency of a firm's production relative to the estimated production frontier, it is important that the frontier is appropriately specified. [12].

Three functional types prevail within the SFA – Cobb-Douglas (CD), CES and translogaritmic (TL) functions. The most general is translog specification of frontier technology as it provides a good first-order approximation to a broad class of functions, including the CES, and includes the Cobb-Douglas as a special case. [12]. CD production function is easy to be estimated (in the linearized form it can be done by ordinary least squares method), and interpreted, because the coefficients are elasticities. The limitations of CD production function are "the inherent assumption of constant elasticity of substitution between the inputs which also implies a constant percentage of income distribution across them." [16]. Unfortunately, the Cobb-Douglas still fits the data well in cases where some fundamental assumptions are violated. [16].

Hence, there are other types of functions used. CD is a special case in a more general class of production functions with constant elasticity of substitution. The translog function relaxes the assumption of constant elasticity of substitution and reduces to the Cobb Douglas function in case there is constant elasticity of substitution. [20].

Several authors compared the usage of both, CD and TL, specifications. For example, Constantine, Martine and Rivera [7] tested the hypothesis that all the second order coefficients and the cross products in TL function are equal to zero. They rejected the null hypothesis, so the specification in the form of a TL function is preferred to the CD. Duffy and Papageorgiou [8] who examined panel data for 82 countries in a period of 28 years also rejected the Cobb-Douglas specification. Similarly, Kneller and Stevens [12] preferred TL specification to CD.

2 Methods

The CES is a natural extension of the Cobb-Douglas in that it permits the elasticity of substitution to be something other than unity – the elasticity of scale can change with output and/or factor proportions. [9] We compare the technical efficiency calculated with Cobb-Douglas production function estimated as (1) and general translog production function estimated as (2). Both functions have the advantage that they can be linearized by natural logarithms. We do not include time t as an explanatory variable, so we cannot assess technical change (change of technology in time), but in this way, the CD and TL models are better comparable.

$$\ln y_{it} = \sum_{k=1}^{K} \beta_k \ln x_{k,it} + \varepsilon_{it} \tag{1}$$

$$\ln y_{it} = \sum_{k=1}^{K} \beta_k \ln x_{k,it} + \frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \beta_{kl} \ln x_{k,it} \ln x_{l,it} + \varepsilon_{it} \tag{2}$$

where y_{it} denotes the production of farm i (i = 1, 2, ..., N, where N is total number of farms) in time t (t = 1, 2, ..., T, where T is total number of observed years). $x_{k,it}$ stands for the input k (k = 1, 2, ..., K, where K is total number of production factors, 4 in our case) of firm i in time t; β_k are the estimated parameters of inputs; ε_{it} is the synthetic error term that consists of inefficiency term u_{it} and stochastic term v_{it}. The division on $\varepsilon_{it} = v_{it} - u_{it}$ was introduced by Aigner et al. [1] who proposed a method that distinguishes the inefficiency from other sources of disturbance that cannot be influenced by the firm. The distribution of the inefficiency was assumed to be truncated normal $u_{it} \sim N^{+}(\mu, \sigma_u^2)$ – with mean μ and variance σ_u^2.

We used True fixed-effects model. It is time-varying model which means that firms' technical inefficiency can develop over time. The parameters of stochastic frontier function are estimated by the maximum likelihood method. Then the elasticities were calculated. "Input elasticities measure the sensitivity of output to an increase in inputs." [2] Coefficients in Cobb-Douglas power function can be already interpreted as elasticities. In translog function are derived by partial derivation of ln y over ln x. The elasticities are calculated at the mean value of the inputs. An example of partial derivation for $x1$ is written in [3].

$$\frac{\delta \ln y}{\delta \ln x_1} = \beta_1 + \beta_{11} \ln x_1 + \frac{1}{2}\beta_{12} \ln x_2 + \frac{1}{2}\beta_{13} \ln x_3 + \frac{1}{2}\beta_{14} \ln x_4 \tag{3}$$

After the estimation of the model, the efficiency is calculated using Jondrow et al. [11] method which measures the contribution of uit to εit $E[u_{it}|\varepsilon_{it}]$. Technical efficiency is then calculated as $\exp[-E(u_{it}|\varepsilon_{it})]$.

Both models were compared by Likelihood ratio test. The differences in technical efficiency were tested by Wilcoxon signed-rank test. In order to understand the effect of specification on efficiency scores, it is also appropriate to examine the correlations between the series by Spearman's rank coefficient ρ. Value of the coefficients takes value between −1 to 1, where values below 0 stays for indirect dependency and above 0 for direct. The calculations were done in Stata IC 15.

Data for the unbalanced panel accounting data of agricultural holdings in the Czech Republic were taken from Albertina database (business register) purchased from Bisnode, s. r. o., company and data about acreage were taken from Land Parcel Identification System. Production of *i*-th farm in time t was approximated by sales of production and services which was adjusted by index of agricultural producers' prices to eliminate the influence of inflation. Explanatory variables – production factors – were: $x_{1,it}$ – consumption of material and energy, $x_{2,it}$ – fixed assets – both variables were adjusted by the index of industry producers' prices in order to eliminate the influence of inflation, $x_{3,it}$ – number of employees and $x_{4,it}$ – acreage of agricultural land. There were 517 farms and 1708 observations for years from 2013 to 2016. Number of observations ranged from 2 to 4 with average 3.3 observations per farm. The description of the variables is given in Tab. 1. Annual sales (58 mil. CZK) were on average higher than the median (40 mil. CZK). Similarly, material and energy consumption was on average 29 mil. CZK with median 21 mil. CZK. Fixed assets were 69 mil. CZK on average, but the median was lower – almost 57 mil. CZK. Average number of employees was 42 with median 38, maximum employees 225.

Variable	Mean	Std. dev.	Median	Min	Max
y_{it} – sales of own products and services (adjusted)	58,453	64,245	40,303	27	498,372
x_{1it} – material and energy consumption (adjusted)	29,057	30,932	21,404	1	315,756
x_{2it} – fixed assets (adjusted)	69,011	90,522	56,698	1	1,200,000
x_{3it} – number of employees [–]	42	42	38	3	225
x_{4it} – acreage [ha]	49,447	1,156	739	0	10,381

Source: own elaboration

Table 1 Description of the variables (in thous. CZK if not stated otherwise)

3 Results

The results of estimated true fixed effects model – with translog and Cobb-Douglas production functions are displayed in table 2. Both models were statistically significant according to Wald χ^2 test. Log likelihood in CD was 1181.8458 and in TL as 1028.4140.

The coefficients in Cobb-Douglas function are interpreted as elasticities. Increase of material and energy consumption by 1% cause increase of production by 0.2803%. Increase of the fixed assets use by 1% cause increase of production by 0.2614%. The highest elasticity is in case of the employees. Increase of their number by 1% cause increase of production by 0.5629%.

The elasticities in translog production function are evaluated at the mean value of the inputs. An increase of material and energy consumption by 1% cause increase of production by 0.3539%. Increase of the fixed assets use by 1% cause increase of production by 0.2627%. Again, the highest elasticity is in case of the employees almost 49.80%. Increase of the acreage of the farm by 1% cause increase of the production by 24.21%. The elasticity is almost similar in case of fixed assets. In case of material and acreage it is higher in translog function. Elasticity of employees is higher in CD function. Sum of the coefficients indicates the returns to scale. In this case it is equal to 1.2085 which shows that firms exhibit increasing returns to scale.

Models were compared using LR test. Model CD is special case of TL model. Likelihood ratio test assumed under H_0 that $\beta_{11} = \beta_{12} = ... = \beta_{44} = 0$, i.e. that production function is better described by the CD function. The likelihood ratio test has test criterion $\chi^2 = 141.95$ that is significant at the 0.05% level. That means that the

probability that H_0 holds is almost zero, hence, the H_0 is rejected. At least one parameter is statistically significantly different from zero. TL is preferred over the CD. Kneller and Stevens [12], Duffy and Papageorgiou [8] and Constantine, Martine and Rivera [7] also rejected the CD specification of production in favor of more general TL.

Then the technical efficiency was estimated using JLMS method [11]. The results are displayed in table 3.

	Cobb-Douglas production function		Translog production function		
ln y	**Coef.** sign. level	**(Std. error)**	**Coef.** sign. level	**(Std. error)**	**Elasticities**
Frontier					
β_1 [ln x_1]	0.2803***	$(8.26 \cdot 10^{-6})$	0.1824***	$(3.46 \cdot 10^{-4})$	0.3537
β_2 [ln x_2]	0.2614***	$(5.44 \cdot 10^{-6})$	0.1744***	$(1.45 \cdot 10^{-4})$	0.2627
β_3 [ln x_3]	0.5629***	$(2.55 \cdot 10^{-5})$	0.1371***	(.)	0.4980
β_4 [ln x_4]	0.1040***	$(3.69 \cdot 10^{-6})$	0.0297***	$(3.07 \cdot 10^{-4})$	0.2421
β_{11} [ln x_1 ln x_1]			0.0088***	$(2.47 \cdot 10^{-5})$	
β_{12} [ln x_1 ln x_2]			0.0030***	$(1.36 \cdot 10^{-5})$	
β_{13} [ln x_1 ln x_3]			0.0205***	$(3.09 \cdot 10^{-4})$	
β_{14} [ln x_1 ln x_4]			0.0048***	$(1.00 \cdot 10^{-5})$	
β_{22} [ln x_2 ln x_2]			0.0060***	$(1.94 \cdot 10^{-5})$	
β_{23} [ln x_2 ln x_3]			0.0220***	$(1.61 \cdot 10^{-4})$	
β_{24} [ln x_2 ln x_4]			0.0039***	$(2.08 \cdot 10^{-5})$	
β_{33} [ln x_3 ln x_3]			0.0249***	$(3.52 \cdot 10^{-4})$	
β_{34} [ln x_3 ln x_4]			0.0074***	$(3.35 \cdot 10^{-5})$	
β_{44} [ln x_4 ln x_4]			0.0141***	$(1.61 \cdot 10^{-5})$	
μ_u					
constant			−14.6404***	(4.5526)	
Variance of inefficiency term					
constant	−3.2042***	(0.0484)	1.0078***	(0.3022)	
Variance of stochastic term					
constant	−32.8654***	(92.6917)	−33.1247	(44.8065)	
σ_u	0.2015***	(0.0049)	1.6552	(0.2501)	
σ_v	$7.30 \cdot 10^{-8}$	$(3.38 \cdot 10^{-6})$	0.0000	$(1.44 \cdot 10^{-6})$	
λ	2759543***	(0.0049)	25800000	(0.2501)	

Source: own elaboration; Note: *** denotes significance at 1%, ** at 5% and * at 10%

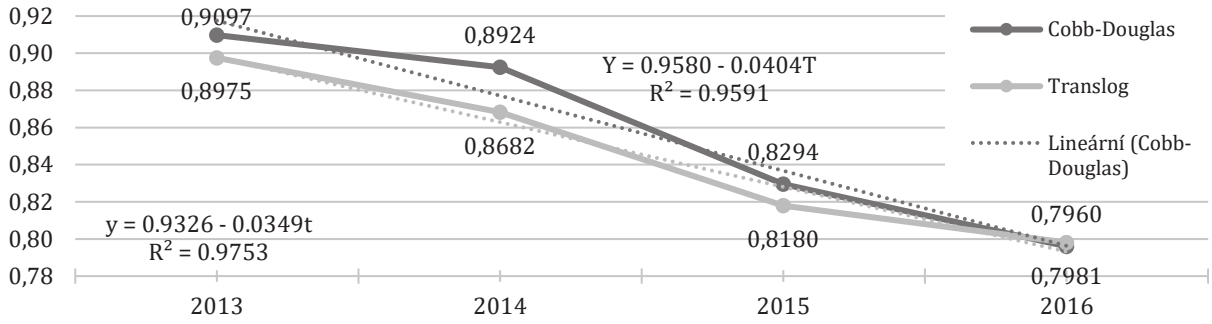
Table 2 True fixed-effects model estimates, distribution of inefficiency term: CD – truncated normal, TL – exponential, distribution of random error: CD, TL – normal

Average technical efficiency in case of CD function was slightly higher (85.69%) than in TL (84.57%). Median was similar in both specifications – 90%. There were 545 100% efficient firms (with efficiency higher than 99%) in CD and 549 in TL. Because the technical efficiency is not normally distributed among farms, the correlation had to be assessed by Spearman correlation coefficient. Spearman's rho was equal to 0.94. H_0 that the efficiencies are independent was rejected. There is positive and strong dependence between both variables.

Production function	Mean	Median	Std. dev.	Min	Max	100% efficient farms
Cobb-Douglas	0.8569	0.8999	0.1668	0.0810	1.0000	545
Translog	0.8457	0.9009	0.1813	0.0693	1.0000	549

Table 3 Technical efficiency of farms based on two specifications of production function (Source: own elaboration)

We can also look on the development of both technical efficiencies in time. From Fig 1, it can be seen that CD efficiency was higher in all years with exception of 2016. It decreased in all years and was the lowest in the last year. This decreasing trend is rather surprising. If we estimate the trend of CD efficiency by trend function, $y = 0{,}9580 - 0{,}0404t$ with rather high fit – adjusted coefficient of determination was $\hat{R}^2 = 93.76\%$. Also, TL trend function $y = 0.9326 - 0.0349t$ had high fit of data – $\hat{R}^2 = 96.29\%$. Average decrease during the whole period 2013 to 2016 was 4.35% in CD function and 3.84% in TL function.



Source: own elaboration

Figure 1 Development of technical efficiency estimated by CD and TL function in period 2013–2016

4 Discussion

Kneller and Stevens [12] concluded "that the effect of the specification of the frontier on measures of productive efficiency is less important." As same as in our case, when TL is preferred over the CD, Kneller and Stevens [12], Duffy and Papageorgiou [8] and Constantine, Martine and Rivera [7] also rejected the Cobb-Douglas specification in favor of translog. The effect of functional form on the estimated efficiency terms was relatively minor. Higher impact had different specification of production function labor and its measurement. Kumbhakar, Heshmati and Hjalmarsson [14] considered the issues related to the specification and estimation of various models incorporating time-varying technical efficiency. They compared Cornwell, Schmidt and Sickles model, Lee and Schmidt model and Battese and Coelli model and found that the efficiency scores vary substantially among the models. "Therefore, an important issue is to what extent the different models generate reasonable, or rather unreasonable, results." [14]. In our case, all models bring reasonable and very similar results. Holmgren [10] compared the efficiency estimated by stochastic cost frontier models with various specification of the explained variable. Particularly he used different expressions of "production of public transport" and concluded that "models using only vehicle-kilometers or only passenger trips tend to underestimate efficiency compared to a model using both at the same time" [10], so also the specification of the variables is important to obtain correct results.

Regarding the height of the technical efficiency in the Czech Republic, it is usually estimated similarly as in our research (around 85%). For example, Pechrová [19] found out that the average efficiency of Czech farms in years 2007–2013 was 86.74%, median was higher (91.03%) as same as in our research. Rudinskaya et al. [21] estimated the technical efficiency of the Czech farms in years 2011–2015 in height of 71.5 %. Lower efficiency was estimated by Nowak, Kijek and Domanska [18] because they used DEA. Under constant returns to scale they calculated average efficiency of Czech agriculture between 2007–2011 in height 69.2% and under variable returns to scale 69.6%. Similarly, in research of Bojnec et al. [3] DEA technical efficiency score of Czech farms for years 2001–2006 varied between 54% in the first year to 65% in the last year of this period.

Regarding the development in time, Čechura et al. [5] found for the CR a trend function y = 0.140 – 0.038t, but with very low fit $\hat{R}^2 = 40\%$. The trend was negative and also average change of technical efficiency during observed period 2004–2011 was negative (−0.029). According to Čechura [4] efficiency in Czech agriculture was considerably volatile in years 2004–2007. Technical efficiency increased in 2005 and decreased in the next year. In 2007, the level of technical efficiency returned to roughly the level it had in 2004. Those two results correspond to our results for the period 2013–2016, so we can see that the decreasing trend of the technical efficiency continued. Naglová and Šimpachová Pechrová [17] found out that food processing companies in the Czech Republic were the least technically efficient in 2020 (average efficiency was 59.21%) and the increase after this year was relatively slow – up to 69.56% in 2014, but then the decrease followed on 65.46% in 2017.

Conclusion

The paper evaluated two types of production function – Cobb-Douglas (CD) and translog (TL) – when estimating the technical efficiency with Stochastic Frontier Analysis. Both specifications were compared on the unbalanced panel data of agricultural holdings in the Czech Republic. Technical efficiencies were calculated by JLMS method and their differences tested by Wilcoxon signed-rank test and correlated by Spearman's rank coefficient.

There were found statistically significant differences in the technical efficiency in CD specification (85.69%) and in TL specification (85.12%), but Spearman's $\rho = 0.9597$ was close to 1 pointing on high positive dependence. Therefore, likelihood ratio test was used to choose the best specification. The CD specification of production function was rejected in favor of more general specification of TL function. We can conclude that TL function is more appropriate to be used as a production function in SFA, but the effect on the technical efficiency is mild and almost negligible despite the statistically significant differences. This is a desirable feature as the specification of the production function does not influence the results of technical efficiency. The challenge for future research is to examine the other determinants of the technical efficiency of the agricultural holdings in the Czech Republic.

Acknowledgements

The paper was supported by the Ministry of Agriculture of the Czech Republic, institutional support MZE-RO0918, Internal Research Project no. 1113/2020.

References

[1] Aigner, D., Lovell, C. & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*, 21–37.

[2] Archambault S. et al. (2014). Nannochloropsis sp. algae for use as biofuel: Analyzing a translog production function using data from multiple sites in the southwestern United States. *Algal Research*, *6*, 124–131.

[3] Bojnec, Š., Fertő, I., Jambor A. & Tóth, J. (2014). Determinants of technical efficiency in agriculture in new EU member states from Central and Eastern Europe. *Acta Oeconomica*, *64(2),*197–217.

[4] Čechura, L. (2012). Technical efficiency and total factor productivity in Czech agriculture. *Agricultural Economics – Czech*, *58(4)*, 147–156.

[5] Čechura, L. et al. (2015). Comparison of Technology and Technical Efficiency in Cereal Production among EU Countries. *Agris on-line Papers in Economics and Informatics*, *VII(2)*, 27–37.

[6] Charnes, A., Cooper, W. & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *3*, 429–444.

[7] Constantin, P. D., Martin, L. M. & Rivera, E. B. B. R. (2009). Cobb-Douglas, Translog Stochastic Production Function and Data Envelopment Analysis in Total Factor Productivity in Brazilian Agribusiness. Journal of Operations and Supply Chain Management, 2(2), 20–34.

[8] Duffy, J.; Papageorgiou, C. (2000). A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth*, *5*, 87–120.

[9] Heathfield, D. F. & Wibe, S. (1987). T*he Translog Function. In: An Introduction to Cost and Production Functions*. Palgrave, London.

[10] Holmgren, J. (2018). The effects of using different output measures in efficiency analysis of public transport operations. *Research in Transportation Business and Management*, 28, 12-22. DOI: 10.1016/j.rtbm.2018.02.006

[11] Jondrow, J., Lovell, C. A. K., Materov, I. S., Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics, 19(1982), 233–238.

[12] Kneller, R. & Stevens, P. A. (2003). The specification of the aggregate production function in the presence of inefficiency. *Economics Letters*, *81*, 223–226.

[13] Kumbhakar, S. C. & Tsionas, E. G. (2006). Estimation of stochastic frontier production functions with input-oriented technical efficiency. *Journal of Econometrics*, *133*, 71–96.

[14] Kumbhakar, S. C., Heshmati, A. & Hjalmarsson, L. (1997). Temporal patterns of technical efficiency: Results from competing models. *International Journal of Industrial Organization*, *15(5)*, 597–616.

[15] Meeusen, W. & Van den Broek, K. (1977). Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review*, *18*, 435–444.

[16] MILLER, E. (2008). An assessment of CES and Cobb-Douglas Production Functions. [on-line] Congres. Budget Office. Available from https://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/94xx/doc-9497/2008-05.pdf.
[17] NAGLOVÁ, Z. & ŠIMPACHOVÁ PECHROVÁ, M. (2019) Subsidies and technical efficiency of Czech food processing industry. *Agricultural Economics*, *65(4)*, 151–159.
[18] NOWAK, A., KIJEK, T. & DOMAŇSKA, K. (2015). Technical efficiency and its determinants in the European Union agriculture. *Agricultural Economics – Czech, 6(16),* 275–283.
[19] PECHROVÁ, M. (2015). Impact of the Rural Development Programme Subsidies on the farms' inefficiency and efficiency. *Agricultural Economics – Czech*, *61(5)*, 197–204.
[20] PRASAD, R. (2015). The production function methodology for estimating the value of spectrum. *Telecommunications Policy*, *39(1)*, 77–88.
[21] RUDINSKAYA, T., HLAVSA, T. & HRUŠKA, M. (2019). Estimation of technical efficiency of Czech farms operating in less favoured areas. *Agricultural Economics – Czech*, *65(10)*, 445–453.
[22] SILVA, T. C. ET AL. (2017). A comparison of DEA and SFA using micro- and macro-level perspectives: Efficiency of Chinese local banks. Physica A: Statistical Mechanics and its Applications, 469, 216–223.