https://doi.org/10.17221/125/2022-AGRICECON

Financial efficiency analysis of Hungarian agriculture, fisheries and forestry sector

Veronika Fenyves[1], Tibor Tarnóczi[1], Zoltán Bács[1], Dóra Kerezsi[1], Péter Bajnai[1], Mihály Szoboszlai[2]*

[1]*Institute of Accounting and Finance, Faculty of Economics and Business, University of Debrecen, Debrecen, Hungary*
[2]*Hungarian National Bank, Hungary*

**Corresponding author: fenyves.veronika@econ.unideb.hu*

Citation: Fenyves V., Tarnóczi T., Bács Z., Kerezsi D., Bajnai P., Szoboszlai M. (2022): Financial efficiency analysis of Hungarian agriculture, fisheries and forestry sector. Agric. Econ. – Czech., 68: 413–426.

Abstract: In this study, we examine the efficiency of companies in Hungary's agriculture, fisheries and forestry sector. We analysed corporate efficiency by using stochastic frontier analysis (SFA). We used two methods to perform the SFA calculations – the Cobb-Douglas and translog functions. The result variable for the SFA calculation was gross value added (GVA), and the explanatory variables were tangibles, material costs, employee costs and other costs. The original database contained cross-sectional and time series data and was transformed into a panel database. We used the maximum log-likelihood method for parameter estimation. We performed the efficiency analysis in the case of the Cobb-Douglas and translog functions in two ways – first, without *z* variables (factor effects) and second, considering different factors (subsectors, workforce categories, ranking by total assets and ranking by total sales). Taking *z* variables into account increased the value of the efficiency coefficients. The latter model's results show that the companies' average performance in the sector examined was more than 70%. Further calculations also showed that the subsectors of the agriculture, fisheries and forestry sector differed in efficiency scores. The larger companies operated more efficiently than the smaller ones in the sector examined.

Keywords: sectoral analysis; performance analysis; production functions; stochastic frontier analysis

The main objective of the research was to examine the efficiency of companies belonging to Hungary's agriculture, fisheries and forestry sector. Because many companies were available, we chose the parametric method of efficiency analysis, stochastic frontier analysis (SFA), to examine efficiency. We examined the efficiency of the selected sector by using the companies' financial statements, considering a specific company size as a lower limit. The aim was also to analyse how the different grouping aspects influenced the evolution of the efficiency scores.

The various sectors of agriculture have been investigated using the SFA method by many authors in Europe and Hungary (Simon and Novák 2002; Gorton and Davidova 2004; Latruffe et al. 2012; Bojnec and Latruffe 2013; Novotná and Volek 2014; Baráth and Fertő 2015; Zsarnóczai and Zéman 2019). As usual in agricultural economics, their studies were based on the Farm Accountancy Data Network (FADN) database. The analysis mostly involved using sales revenue as an independent variable.

The research we present in this study is not based on the FADN database but on company financial reports, balance sheets and income statements submitted to the tax authority. We used gross value added (GVA) as an independent variable. The use of GVA

Supported by the Hungarian National Bank.

https://doi.org/10.17221/125/2022-AGRICECON

has one disadvantage compared with sales revenue: its value can be negative. The regression functions used in the SFA method are mostly logarithmic, meaning that companies with negative and 0 values are not included in the study.

The firms' GVA often measures economic performance, which is a crucial business performance evaluation indicator (Horváthová et al. 2014). Van Passel et al. (2007), Thomassen et al. (2009), and Giannakis and Bruggeman (2015) all used GVA indicators to measure the performance of agricultural firms.

Productivity can be the main indicator to measure performance because a firm usually uses more input to determine the total factor productivity, defined as a ratio of output to total input (Palia and Lichtenberg 1999). Several researchers have used total factor productivity (Kim and Maximovic 1990; Lichtenberg and Siegel 1990; Caves and Barton 1991).

According to Hall and Jones (1999), with data on output and input factors, the productivity level can be calculated directly from the production function. The production function sets the highest possible output limit that a firm can obtain by a combination of factors at the level of technical knowledge over the production period. The maximum output concerns a firm and all other firms in the same industry. Therefore, the function defined has been called an industry production function (Aigner and Chu 1968). Schmidt (1985) defines the production function as 'a function giving the maximum possible quantity of some output, given quantities of a set of inputs'.

Farrell (1957) published a ground-breaking work that showed the possibility of estimating the frontier production functions. Pitt and Lee (1981) defined the production frontier as 'the locus of technically efficient input-output combinations'. On the basis of Farrell's work (Farrell 1957), several authors, such as Aigner and Chu (1968) and Richmond (1974), developed the production frontier estimation later.

Aigner et al. (1977) and Meeusen and van den Broeck (1977) proposed a stochastic frontier production function independently, and it was a significant contribution to the firms' technical efficiency (TE) estimation. The stochastic frontier production function includes two random components – technical inefficiency and random error (Battese and Coelli 1992).

We selected the Cobb-Douglas and translog production functions for this research. The Cobb-Douglas function must satisfy many more conditions than does the translog function; however, the Cobb-Douglas function is more accessible to interpret than is the translog function. The goal was not to analyse the parameters of the obtained functions but to analyse the efficiency of the investigated companies, the sector and its various groups.

According to Kalirajan and Shand (1999), TE should be measured for two main reasons. First, a noticeable gap exists between empirical reality and the theoretical assumption of total TE. Second, it is highly likely that where technical inefficiency exists, it influences allocative efficiency and has a cumulative negative effect on economic efficiency. Therefore, TE measurement is essential for firms to achieve high economic performance (Kalirajan and Shand 1999).

Diaz and Sánchez (2008) explained that companies behave optimally when their production process is technically efficient. The frontier production function, with the maximum output a firm can achieve, represents this behaviour with input and technology.

Baráth and Fertő (2015) used the SFA method to analyse TE and found that technological heterogeneity plays an essential role in Hungarian farms producing cereal, oilseed and protein crops). They showed that the Hungarian cereal, oilseed and protein crop sector has less chance of improving TE performance than expected previously. It is impossible to improve productivity by increasing farm size unless those farms change technologies.

Náglová and Šimpachová Pechrová (2019) used SFA to examine Czech food industry companies' TE. They found that all independent variables were statistically significant. The average TE coefficient was 65.64%, and the TE was statistically significantly different in some regions.

Čechura et al. (2022) used the SFA method to analyse the relationship between farm size and productivity in Czech cereals, milk and beef production. They found that large farms did better than small ones in productivity and TE for the data from 2014 to 2018.

Mitsopoulos et al. (2021) examined the performance of dairy farms in Greece. They identified factors affecting the profitability of farms. They assessed the viability of these farms according to TE.

Various authors in several periods examined the efficiency of Hungarian agricultural companies in Hungary, including Bakucs et al. (2010, 2020), Baráth et al. (2020, 2021), and Kovacs and Szucs (2020). These analyses were mainly based on the FADN database. The FADN database is a so-called test farm database, which, for example, contained 1 753 individual farms and 415 partnerships in 2020 (Keszthelyi 2021). The database is based on samples representing the primary population and

https://doi.org/10.17221/125/2022-AGRICECON

is created by weighting the sample farms, where the weight expresses how many farms the sample represents in the given group of the base population.

MATERIAL AND METHODS

Our analysis data come from the database provided by the Hungarian taxation authority made available by the Hungarian National Bank for joint research. The entire database included the data of all companies that operated under the accounting law. The companies selected for the analysis are in the agriculture, fisheries and forestry sector (Sector A) in Hungary according to the Hungarian unified sectoral classification system of economic activities. We excluded companies from the analysis if their financial statements were incomplete or if they had no financial statements for all years examined (2017–2019). We also excluded companies that had incorrect data in the analysed variables (for example, negative material cost, employee expenses or depreciation) and the companies with less than 30 000 EUR in total revenues and total assets in their financial statements. After the exclusions, 4 034 firms remained in the database yearly, including 12 102 records.

Sector A has a small proportion of the total GVA, as shown in Table 1. Although several companies were excluded from the database, Table 1 shows that the GVA did not differ on average from the value listed by the Hungarian Central Statistical Office. This finding means that the results can provide satisfactory conclusions for the whole sector. The total GVA of the database companies examined exceeded the total GVA of Sector A because the excluded firms mostly had negative values. We also classified the data in the database according to subsectors and workforce number (Tables 2, 3).

Table 2 shows that more than 40% of companies employed only 1 to 4 people, and approximately 90% had between 1 and 49 people. Furthermore, 3% of companies did not have any employees, and only 6% of companies employed more than 50 people. It is important

Table 1. The proportion of sector gross value added (GVA) in total economy GVA (%)

Year	100% = total economy	100% = Sector A
2017	2.77	120.69
2018	2.57	89.76
2019	2.51	93.01
Average	2.61	100.89

Source: Own calculation from online data of Hungarian Statistical Office (https://statinfo.ksh.hu/Statinfo/haDetails.jsp)

Table 2. Distribution of companies by year and workforce number category

Category number	Persons in category	Number of companies 2017	2018	2019
1	0	128	124	120
2	1–4	1 690	1 692	1 684
3	5–9	894	892	891
4	10–19	643	656	666
5	20–49	422	421	425
6	150–249	233	226	226
7	≥ 250	24	23	22

Source: Own calculation from the database received from the Hungarian Tax Authorities.

to note that companies with 0 employees are also active enterprises, of which there are quite a few examples in Hungary. There are no significant differences between the years regarding the employee number. Thus, 93% to 94% of companies were in the small and medium-sized enterprises (SMEs) category regarding employee numbers.

Table 3 shows that more than one-half of the analysed companies were engaged in crop production as their main activity (56.3%) according to the Hungarian unified sectoral classification system of economic activities code they provided. In addition, almost 21% of them were engaged in animal husbandry. Agricultural services accounted for the third-largest share (11.43%).

We categorised the companies into two groups according to size. We created the ranking based on the yearly average values so that a given company was placed in the same class every year. First, we split the total assets and total revenue variables into deciles. Next, we classified the values of the two variables into ten groups. Furthermore, we used the two ranks developed to determine whether they significantly affected the SFA model's efficiency coefficients (ECs).

Before using the SFA method, we performed multivariate logarithmic regressions with the same variables to test multicollinearity. We used the variance inflation factor (*VIF*) function of the R statistics' car package to test multicollinearity with the *VIF*. According to the literature, collinearity causes problems when the *VIF* value exceeds 5 (Sheather 2009; Petrie 2016). From the formula calculating the *VIF*, it is possible to define critical *VIF* values, for which we used the following formula:

$$VIF_{crit} = \frac{1}{1 - R^2} \quad (1)$$

where: R^2 – coefficient of determination.

https://doi.org/10.17221/125/2022-AGRICECON

Table 3. Distribution of companies by TEÁOR code

Category code	Subsector name	Number of companies	Distribution of companies (%)
11	Production of non-perennial crops	1 952	48.39
12	Production of perennial crops	258	6.40
13	Production of plant reproductive material	61	1.51
14	Animal husbandry	842	20.87
15	Mixed farming	60	1.49
16	Agricultural services	461	11.43
17	Wildlife management	15	0.37
21	Forestry activities	121	3.00
22	Wood production	95	2.35
23	Collecting wild forest products	2	0.05
24	Forestry services	105	2.60
31	Fishing	15	0.37
32	Fish management	47	1.17
Total		4 034	100.00

Source: Own calculation from the database received from the Hungarian Tax Authorities

The multivariate regression calculation also considered the effect of the year by including the year variable.

We chose SFA for our research. Several packages in the R statistical system perform SFA-related calculations, and we selected the frontier package for this research.

We used the TE of the output side to measure efficiency and the company's ability to achieve maximal output from inputs. In microeconomics, the production function assigns the maximal output for each input – that is, the technically efficient output level. Therefore, the observations cannot be greater than the production function, and being lower than the production function shows technical inefficiency. Therefore, all residuals have to be negative or 0 values:

$$\ln(y) = \ln f(x) - u \tag{2}$$

where: u – technical inefficiency ($u \geq 0$).

Statistical noise, such as measurement errors, variables omitted from the function and estimation errors, exists in all databases and models. The stochastic frontier model includes both technical inefficiency (u) and statistical noise (v), and Equation (2) is modified this way:

$$\ln(y) = \ln f(x) - u + v \tag{3}$$

where: v – statistical noise (Coelli et al. 2005).

Equation (3) can be transformed as follows:

$$y = f(x) \times e^{-u} \times e^{v} \tag{4}$$

where: e^{-u} – value of technical inefficiency; e^{v} – value of statistical noise.

Output-oriented *TE* can be defined as the ratio of observed and marginal output (Coelli et al. 2005):

$$TE = \frac{y}{f(x)e^{v}} = \frac{f(x)e^{-u}e^{v}}{f(x)e^{v}} = e^{-u} \tag{5}$$

We used the Cobb-Douglas and translog functions for the regression function, and we used a maximum log-likelihood estimation to determine the coefficients of the stochastic frontier model. We used the SFA function of the frontier package to determine the coefficients of the regression functions and the efficiency and inefficiency values. The SFA function defines additional values besides the regression parameters to help evaluate the results. One such value is γ, which is the ratio of the variance of the inefficient part (u) to the total variance (ε). If γ is 0, then u is irrelevant, and the result is the same as in the ordinary least squares (OLS) model. However, if it is 1, then v is irrelevant. Statistical noise (v) and inefficiency (u) are essential to explain deviations from the production function, but inefficiency is more important than noise. The absence of ineffi-

https://doi.org/10.17221/125/2022-AGRICECON

ciency can be checked by using the likelihood ratio test, where the null hypothesis is that $\gamma = 0$. The test result indicates whether the u value significantly improves the model's fit. The likelihood ratio test compares the log-likelihood values of two models.

The other important ratio is λ, which is the ratio of the standard deviation of u to v. The third ratio is the inefficiency index (*InEff*), which shows the proportion of inefficiency variance to the total variance:

$$InEff = \frac{\sigma_u^2}{\left(\sigma_u^2 + \sigma_v^2\right)} \quad (6)$$

where: σ_u^2, σ_v^2 – variance of technical inefficiency and statistical noise, respectively.

If a firm has no inefficiencies ($u = 0$) for all observations, then the γ indicator would also be 0, so the null hypothesis of inefficiency can be tested in terms of whether γ is not significantly different from 0. We can check this with a likelihood ratio test to see whether including the u inefficiency term significantly improves the model's fit. The frontier package provides the lrtest function to perform this test. If the lrtest function is called with only one stochastic frontier model, it compares the SFA model with the corresponding OLS model. This comparison is possible because the SFA function of the frontier package also calculates both the log-likelihood and the OLS model. If the probability value of lrtest is less than 0.05, then the OLS model can be rejected compared with the stochastic model – that is, there are significant technical inefficiencies.

The SFA calculation assumes that the investigated companies use similar technology for their activities, which is not valid for a more complex database. Unobservable differences between technologies may be incorrectly detected as inefficiencies if technological differences are not considered. Accordingly, it is advisable to classify the observations into different classes. The created groupings can be used in several ways – that is, the heterogeneity problem can be solved in several ways. A two-phase calculation is possible, in which the production function is first calculated with the Cobb-Douglas or translog equation, and then another regression calculation is performed with calculated u values as the dependent variable (Battese and Coelli 1995):

$$u = \delta_0 + \delta_1(class_1) + \delta_2(class_2) + \ldots + \delta_m(class_m) + W_{it} \quad (7)$$

where: u – technical inefficiency; δ_i – indicates coefficients of the regression function; $class_i$ – indicates various factors that indicate the supposed differences between companies; W_{it} – indicates a random variable defined by the truncated normal distribution with 0 mean and σ_2 variance.

Abdul et al. (2022) used the previous model to verify the determinants of technical inefficiency. The factors used were the farmers' education level, age, planting system, seed quality, pests, extension services, credit, membership of a farmer association and being a plasma farmer. Abate et al. (2022) also used the same model to analyse the efficiency of white cumin production in northwestern Ethiopia. The difference between the two studies is that Abdul et al. (2022) used a separate model, whereas Abate et al. (2022) treated factors affecting inefficiency as a combined model.

We uploaded the latest version (1.1-8) of the R statistics frontier package used for the calculations on April 17, 2020. Henningsen (2020) used Coelli's FRONTIER 4.1 program to create this program; however, this conference paper also included a three-step approach to calculating efficiency.

In the R frontier package, the z variables can be entered together with the SFA function, but it means the same as what was presented in Equation (7). In this case, specifying the regression function consists of two parts:

$$y = f(x)|f(z) - u + v \quad (8)$$

Both $f(x)$ (endogenous variables) and $f(z)$ (exogenous variables) symbolise a regression relationship with the regression constant and coefficients.

The second option is to use the so-called latent variable SFA method, which combines SFA and latent variable models. In a latent class model, the unconditional likelihood value of the i^{th} company is obtained as the weighted sum of the j-class likelihood functions, where the weights are the class membership probabilities (Orea and Kumbhakar 2004). This method provides a new aspect to considering the technology effect. We did not present the latent class model in this study. The z variables are also suitable from a certain point of view to consider the model's heterogeneity. The latent class model in the sfaR package in R statistics provides the possibility for complex analysis.

The analysis database includes all companies subject to the Hungarian Accounting Act that have filed a tax return. At the same time, the analysis database also contains certain narrowings compared with the entire database, which is not significant.

https://doi.org/10.17221/125/2022-AGRICECON

In this research, we considered the *GVA* as a corporate value creation indicator. This indicator is beneficial because it ignores financing and taxation effects, varying from company to company. Furthermore, for this research, we used the *GVA* as a result variable and examined a given sector of the national economy, which had not been analysed in Hungary. It can also be helpful for foreign researchers to know the results of such research, both from the point of view of the specific sector and the methods used.

In this study, we examined the performance of Hungarian agricultural companies for the period from 2017 to 2019. Considering the literature, the methodology and the objective of the research, we formulated the hypotheses as follows:

1) The groups differ statistically significantly according to the examined attributes and the created groupings.
2) Considering the created classes improves efficiency scores.
3) The performance of companies in Sector A is relatively good.
4) There are significant differences among the average performances of sectors.
5) The efficiency scores of companies decreases with a higher number of employees.
6) A larger value of an asset portfolio does not go hand in hand with increased corporate efficiency.
7) As sales revenue increases, so does company efficiency.

RESULTS AND DISCUSSION

We used multivariate analysis of variance to determine the effect of the subsector, workforce category, total assets and total revenue rankings factors on the variables of the SFA model (GVA, tangibles, material costs, employee costs and other costs). Results from the multivariate analysis of variance Pillai trace showed at least a 0.1% significance level for each factor (Table 4), so there were significant differences among the groups in terms of the variables. Because there were significant differences between the factor groups, we also performed a one-way analysis of variance. Table 4 shows that all examined factors statistically significantly affected GVA, tangibles and other costs at a significance level of at least 1%. Therefore, the variables differed significantly according to the examined factors. However, the subsector factor did not affect material costs significantly, and employee costs were not affected by rank by total assets. Therefore, on the basis of the results of the analysis of variance, it is necessary to consider these factors in the efficiency analysis. On the basis of these results, hypothesis 1 is confirmed, except in 2 cases out of the 24, because the groups based on each factor were significantly different.

Because we used the logarithmic model for the SFA calculation, we performed multivariate logarithmic regression with the variables of the SFA function to determine the relationship between dependent and independent variables and test the multicollinearity. However, the number of data (for 2017, 3 889; for 2018, 3 866; for 2019, 3 890) in this model differs from the original number of data (4 034 firms) because companies with negative or 0 values were excluded from the estimation because of the logarithmic transformation. The total decrease in three years was 370 companies.

In the logarithmic model, all regression coefficients were significant at a level of at least 1% (Table 5). In Table 5, the multicollinearity test showed that none of the VIF values exceeded the calculated critical value.

Before using the SFA method, we converted the data to panel data by using the pdata.frame function of the R statistics plm package. We used a true fixed effects model. Thus, the model can consider time series data, and the three years can be analysed together. The results

Table 4. Significance levels of multivariate and multifactorial analysis of variance

Variables	Sub-sector	Workforce category	Ranking by total assets	Ranking by zotal revenue
All variables	***	***	***	***
Gross value added (GVA)	***	***	***	***
Tangibles	***	**	***	***
Material costs	***	–	***	***
Employee costs	***	***	–	***
Other costs	***	**	**	***

**, ***Statistical significance at 1% and 0.1% levels, respectively

Source: Own calculation

https://doi.org/10.17221/125/2022-AGRICECON

Table 5. Results of multivariate logarithmic regression and multicollinearity testing (dependent variable: *GVA*)

Variable names	Regression coefficients	Significance levels (%)	Significance level signs	*VIF* values
Intercept	345.692	0.22	**	–
Tangibles	0.215	0.00	***	2.11
Material cost	0.122	0.00	***	1.60
Employee cost	0.298	0.00	***	2.14
Other costs	0.176	0.00	***	2.25
Year	–45.084	0.24	**	1.00
R-squared		0.779		
Adjusted *R*-squared		0.779		
Critical *VIF* value		4.520		

**, ***Statistical significance at 1% and 0.1% levels, respectively; *VIF* – variance inflation factor; *GVA* – gross value added
Source: Own editing using R statistics results

of the SFA calculation are shown in Table 6, and these results show that all regression coefficients were significant at less than 0.01%. Because the coefficients can also be considered elasticity ratios, Table 6 also shows that the most substantial effect on the development of *GVA* was employee costs, and the smallest was material costs. The variance of the regression residuals was 45.97% of the variance of the dependent variable (1.9295). The inefficiency was 73.64% of the total residual variance (γ), and the noise was 26.36%. These values show that the noise was not irrelevant, but the TE mainly explains the deviation from the production frontier.

Besides the likelihood model, the frontier function also calculates the OLS model. The likelihood ratio test results showed that the probability value for the χ^2 test was less than 0.1%, so the likelihood model produced better results than did the OLS model. This result can also be interpreted as indicating that it is worth splitting the residual variance and determining the inefficiency and noise effects separately.

The results of the SFA model, including all four factors, are also shown in Table 6. The regression coefficients changed after the comparison of the results against the model without *z* variables. The highest regression coefficient decrease was in the coefficient of material costs (27.59%), and the smallest was in the coefficient of employee costs (2.15%). The intercept (7.21%) and the coefficient of other costs (5.04%) increased. These changes mean that the effect of the variables also changed. The coefficient of each independent variable was positive, meaning that each variable value increased the *GVA*. All regression coefficients were significant at 0.1% or greater. Other parameters of the model also changed significantly. The variance of the residuals increased by 140%, and the inefficiency variance increased by 34%. The inefficiency ratio reached 88.08% compared with 65.62% in the nonfactor model. The λ value increased from 1.6735 to 2.7195, and the γ value increased by 20%. We can conclude that considering these factors produced better results than would be possible without them.

We determined the ECs by using the translog function, a quadratic function. The applicability of the translog function is much more flexible than that of the Cobb-Douglas function. The translog function contained the same dependent and independent variables as did the Cobb-Douglas function. The results of the translog function are shown in Table 7. The translog function was also significantly different from the OLS function, so dividing the regression's error term into two parts makes sense. On the basis of the results shown in Table 7, we can conclude that three coefficients were not significant – tangibles, the square of other costs and the interaction between tangibles and other costs. Comparison of Tables 6 and 7 shows that the Cobb-Douglas and translog functions without *z* variables did not differ significantly from each other in terms of parameters related to efficiency scores. The γ, λ and inefficiency ratios were slightly higher in the translog function than in the Cobb-Douglas function. At the same time, the values of the ECs, in comparison to the overall averages, were almost 8.5% higher (0.5808 and 0.6297).

The differences between the Cobb-Douglas and translog functions using the *z* variable were more significant than were those without the *z* variable, yet the averages of the efficiency scores did not differ significantly. The difference between the total averages did not reach 2%, but at the same time, the total average value of efficiency scores was higher for the translog function. The translog function cannot be justifiable be-

https://doi.org/10.17221/125/2022-AGRICECON

Table 6. Results of Cobb-Douglas model using maximum log-likelihood SFA method (dependent variable: *GVA*)

Variables and other parameters	SFA function without *z* variables		SFA function with all *z* variables	
	coefficients	significance level (%)	coefficients	significance level (%)
Intercept	3.8263	0.00***	4.1021	0.00***
ln(*intangibles*)	0.2158	0.00***	0.1985	0.00***
ln(*material costs*)	0.1022	0.00***	0.0740	0.00***
ln(*employee costs*)	0.2656	0.00***	0.2599	0.00***
ln(*other costs*)	0.1706	0.00***	0.1792	0.00***
z_Sector	–	–	0.0504	0.00***
z_EmpCat	–	–	–0.2800	1.49*
z_Rang.TA	–	–	–0.1011	0.22**
z_Rang.Sales	–	–	–0.8076	0.00***
σ^2	0.8863	0.00***	2.1309	0.00***
$\sigma(u)^2$	0.6531	0.00***	1.8770	0.00***
$\sigma(u)$	0.8081	0.00***	1.3701	0.00***
$\sigma(v)^2$	0.2332	0.00***	0.2538	0.00***
$\sigma(v)$	0.4829	0.00***	0.5038	0.00***
γ	0.7369	0.00***	0.8809	0.00***
λ	1.6735	0.00***	2.7195	0.00***
Var(u)	0.2373	–	0.6821	–
Var(v)	0.2332	–	0.2538	–
Inefficiency impact	0.5044	–	0.7288	–
Number of companies	4 005	–	–	–
Number of years	3	–	–	–
Excluded companies	370	–	–	–
Number of items included in the regression	11 645	–	–	–
Total average efficiency	0.5808	–	0.7190	–
Average efficiency 2017	0.5898	–	0.7214	–
Average efficiency 2018	0.5815	–	0.7187	–
Average efficiency 2019	0.5711	–	0.7171	–

*, **, ***Statistical significance at 5%, 1%, and 0.1% levels, respectively; SFA – stochastic frontier analysis; *GVA* – gross value added; *z_Sector* – the sub-sectors within the national economic sector as influencing factors; *z_EmpCat* – workforce categories as influencing factors; *z_Rang.TA* – decile classes (size categories) formed based on the values of all assets as influencing factors; *z_Rang.Sales* – decile classes (size categories) formed based on total sales values as influencing factors σ^2 – total variance of regression error; $\sigma(u)$ and $\sigma(u)^2$ – standard deviaton and variance of the inefficiency part; $\sigma(v)$, $\sigma(v)^2$ – standard deviation and variance of the statistical noise part; γ – proportion of inefficiency variance in total variance; λ – the ratio of the standard deviation of inefficiency and statistical noise

Source: Own editing using the calculation results

cause of this value difference, but it is more appropriate to use it because of its flexibility.

Figure 1 contains the distributions of corporate efficiency scores according to the averages of the years. Figure 1 shows that the efficiency scores for the translog function were higher than those for the Cobb-Douglas function. The differences were larger in the case without the *z* variables. For example, for the translog function, the number of companies with an efficiency score lower than 0.5 decreased by 29.81%, and those between 0.7 and 0.9 increased by 29.51% compared with those for the Cobb-Douglas function. At the same time, the number of companies with an efficiency score greater than 0.9 decreased by 6.74%. The total average of the translog function was 6.3% higher than that of the Cobb-Douglas function. Similar changes can be seen in the functions including *z* variables, but the rate of change was lower (1.86%). For the translog function with *z* variables, the number of companies with an efficiency score lower than 0.5 decreased by 12.11%,

https://doi.org/10.17221/125/2022-AGRICECON

Table 7. Results of the translog SFA model (dependent variable: *GVA*)

Variables and other parameters	SFA function without *z* variables		SFA function with all *z* variables	
	coefficients	significance level (%)	coefficients	significance level (%)
Intercept	5.8690	0.00***	7.1417	0.00***
log(*intangibles*)	–0.0366	34.54	–0.0723	2.70*
log(*material costs*)	–0.1900	0.00***	–0.1785	0.00***
log(*employee costs*)	0.1365	0.23**	–0.0859	2.12*
ln(*other costs*)	0.3670	0.00***	0.4229	0.00***
1/2 × log(*intangibles*)^2	0.0684	0.00***	0.0625	0.00***
1/2 × log(*material costs*)^2	0.0836	0.00***	0.0558	0.00***
1/2 × log(*employee costs*)^2	0.1110	0.00***	0.1294	0.00***
1/2 × log(*other costs*)^2	0.0041	19.56	0.0078	1.62*
log(*intangibles*) × log(*material costs*)	–0.0144	0.00***	–0.0024	43.36
log(*intangibles*) × log(*employee costs*)	–0.0332	0.00***	–0.0365	0.00***
log(*intangibles*) × log(*other costs*)	0.0012	65.36	–0.0022	39.46
log(*material costs*) × log(*employee costs*)	–0.0323	0.00***	–0.0223	0.00***
log(*material costs*) × log(*other costs*)	–0.0143	0.00***	–0.0107	0.03***
log(*employee costs*) × log(*other costs*)	–0.0170	0.00***	–0.0244	0.00***
z_Sector	–	–	0.1524	1.44*
z_EmpCat	–	–	–5.0773	0.05***
z_Rang.TA	–	–	–1.0061	0.05***
z_Rang.Sales	–	–	–1.6497	0.00***
σ^2	0.7443	0.00***	9.5288	0.01***
$\sigma(u)^2$	0.5461	0.00***	9.3414	0.01***
$\sigma(u)$	0.7390	0.00***	3.0564	0.00***
$\sigma(v)^2$	0.1982	0.00***	0.1874	0.00***
$\sigma(v)$	0.4452	0.00***	0.4329	0.00***
γ	0.7337	0.00***	0.9803	0.00***
λ	1.6597	0.00***	7.0607	0.00***
Var(u)	0.1984	–	3.3945	–
Var(v)	0.1982	–	0.1874	–
Inefficiency impact	0.5002	–	0.9477	–
Number of companies	4 005	–	–	–
Number of years	3	–	–	–
Excluded companies	370	–	–	–
Number of items included in the regression	11 645	–	–	–
Total average efficiency	0.6297	–	0.7324	–
Average efficiency 2017	0.6181	–	0.7365	–
Average efficiency 2018	0.6042	–	0.7307	–
Average efficiency 2019	0.6174	–	0.7299	–

*, **, ***Statistical significance at 5%, 1%, and 0.1% levels, respectively; SFA – stochastic frontier analysis; *GVA* – gross value added; ; *z_Sector* – the sub-sectors within the national economic sector as influencing factors; *z_EmpCat* – workforce categories as influencing factors; *z_Rang.TA* – decile classes (size categories) formed based on the values of all assets as influencing factors; *z_Rang.Sales* – decile classes (size categories) formed based on total sales values as influencing factors; σ^2 – total variance of regression error; $\sigma(u)$, $\sigma(u)^2$ – standard deviaton and variance of the inefficiency part; $\sigma(v)$, $\sigma(v)^2$ – standard deviation and variance of the statistical noise part; γ – proportion of inefficiency variance in total variance; λ – the ratio of the standard deviation of inefficiency and statistical noise

Source: Own editing based on calculation results

https://doi.org/10.17221/125/2022-AGRICECON

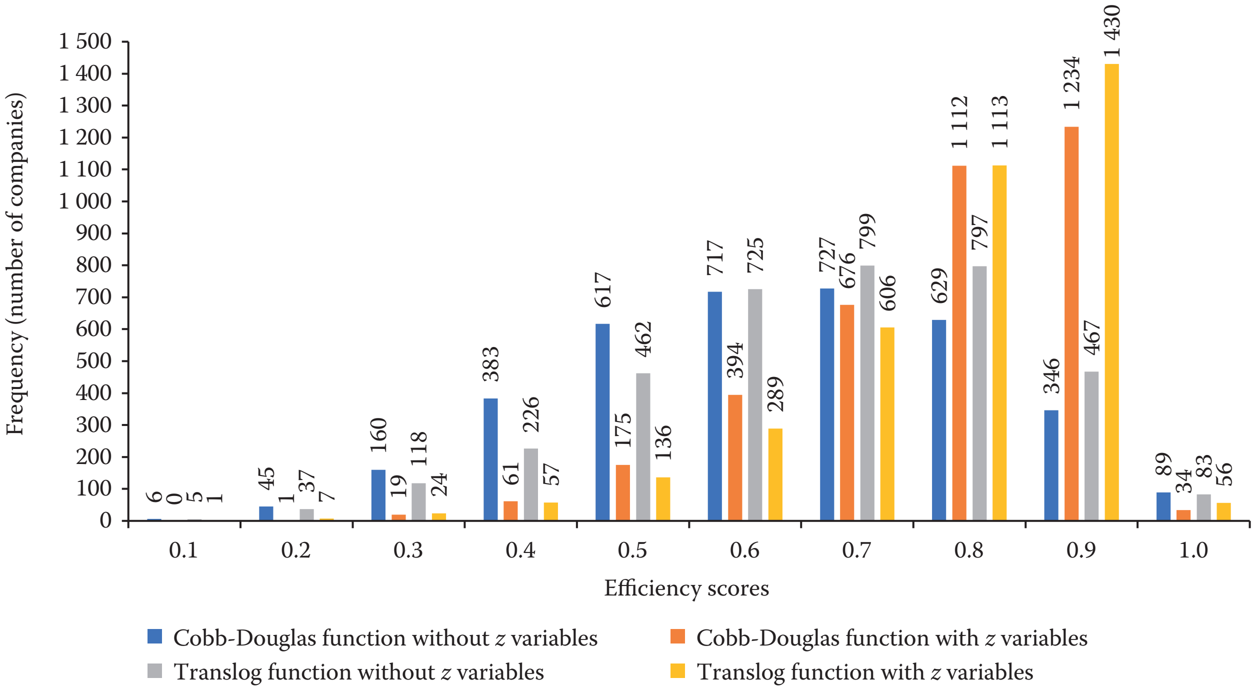

Figure 1. Distribution of corporate efficiency scores considering averages of the years

Source: Authors' elaboration using the results of stochastic frontier analysis

and those between 0.5 and 0.7 decreased by 16.07%. The decrease in the two lower intervals caused an increase in the intervals of 0.7 to 0.9 and greater than 0.9 (8.31% and 17.02%). On the basis of the data in Figure 1, the efficiency scores for the functions, including the *z* variables, were higher, with an average of 23.79% for the Cobb-Douglas function and 18.63% for the translog function. Therefore, there was no significant difference in efficiency scores between the translog and the Cobb-Douglas functions with *z* variables. The use of the translog function cannot be argued because of its more flexible applicability.

Figure 2 shows the kernel density of efficiency scores. Figure 2 was created to illustrate the relative distribution of the efficiency scores determined by the four functions. Figure 2 also supports what was described in Figure 1.

Hypothesis 2 is accepted because the SFA models with *z* variables had higher efficiency scores than did the SFA models without *z* variables. On the basis of these results, hypothesis 3 is also acceptable. The models with *z* variables could be considered as plausible because they exceeded the 0.7 value on average. The distribution of efficiency scores was also favourable.

We calculated average values per sector by using the efficiency scores of the *z* variable Cobb-Douglas and translog models (Table 8). We also assigned ranking numbers to the individual values for both functions. The ranking numbers of the two functions were the same in six cases and different in seven cases. The efficiency score averages did not differ significantly for the two models. The largest difference was 4.29% (code: 23, see Table 8), and the smallest was 0.41% (code: 13, see Table 8). For the Cobb-Douglas function, the difference between the smallest and the largest score was 12.67%, and for the translog function, that difference was 10.64%. Table 8 shows that the production of perennial

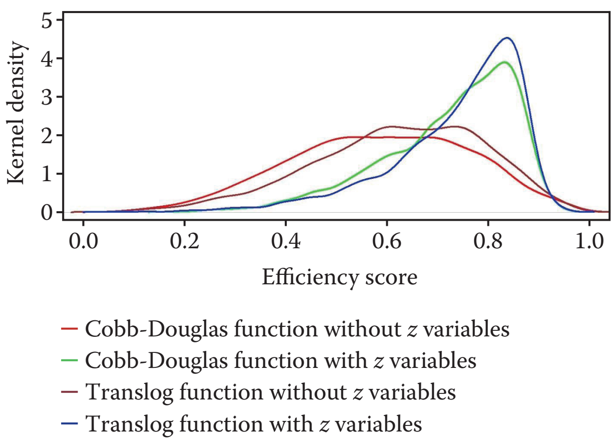

Figure 2. Kernel density of efficiency scores of different stochastic frontier analysis (SFA) functions

Source: Authors' elaboration using the results of stochastic frontier analysis

https://doi.org/10.17221/125/2022-AGRICECON

Table 8. Average efficiency coefficients (ECs) per sector for the models with *z* variables

Category code	Category name	Cobb-Douglas function	Ranking	Translog function	Ranking
11	Production of non-perennial crops	0.7206	4	0.7333	4
12	Production of perennial crops	0.7241	1	0.7415	1
13	Production of plant reproductive material	0.7223	3	0.7252	7
14	Animal husbandry	0.7229	2	0.7372	2
15	Mixed farming	0.7183	6	0.7298	5
16	Agricultural services	0.7155	7	0.7269	6
17	Wildlife management	0.7185	5	0.7238	8
21	Forestry activities	0.7025	10	0.7167	12
22	Wood production	0.6979	12	0.7174	10
23	Collecting wild forest products	0.6427	13	0.6703	13
24	Forestry services	0.7038	9	0.7177	9
31	Fishing	0.6992	11	0.7168	11
32	Fish management	0.7155	7	0.7362	3

Source: Own editing from the R statistics calculations

crops sector had the highest average scores, followed by animal husbandry. The lowest average scores were in the collecting wild forest products sector. These rankings were the same for both models. Because there were detectable differences among the sectors, hypothesis 4 is also accepted.

Table 9 contains the average efficiency scores per workforce category, which in this case was also calculated as an average over the years. Table 9 shows that enterprises with a larger number of employees had a higher efficiency index than did those with a smaller number of employees. Therefore, one could assume that enterprises with a smaller number of employees require more attention. At the same time, this assumption was not verified in the examined sector. Therefore, hypothesis 5 is rejected.

Table 9. Average efficiency scores per workforce category for the models with *z* variables

Workforce category	Cobb-Douglas function	Translog function
0	0.7088	0.7223
1–4	0.7097	0.7216
5–9	0.7164	0.7321
10–19	0.7284	0.7447
20–49	0.7352	0.7476
150–249	0.7423	0.7514
≥ 250	0.7563	0.7599

Source: Own editing from the R statistics calculations

Table 10 shows the change in efficiency scores according to the total asset portfolio's growth. On the basis of the data in Table 10, we can conclude that, in terms of tendency, companies' average efficiency increased with the total asset portfolio's growth, but there were decreases in some categories. The translog function had a higher value than did the Cobb-Douglas function in all categories, and the difference varied between 1.22% and 2.35%, which is not a significant difference. For the Cobb--Douglas function, except for the first rank number, the efficiency score of all rank numbers was greater than 0.7. For the translog function, the efficiency score of each category was greater

Table 10. Average efficiency scores for the models with *z* variables, ranking by total assets

Rank numbers	Cobb-Douglas function	Translog function
1	0.6945	0.7105
2	0.7027	0.7150
3	0.7173	0.7289
4	0.7150	0.7278
5	0.7200	0.7350
6	0.7201	0.7370
7	0.7290	0.7411
8	0.7195	0.7327
9	0.7380	0.7524
10	0.7335	0.7425

Source: Own editing from the R statistics calculations

https://doi.org/10.17221/125/2022-AGRICECON

Table 11. Average efficiency scores for the models with *z* variables, ranking by total revenues

Rank numbers	Cobb-Douglas function	Translog function
1	0.5649	0.5714
2	0.6342	0.6328
3	0.6810	0.6762
4	0.7211	0.7161
5	0.7583	0.7532
6	0.7883	0.7838
7	0.8094	0.8106
8	0.8335	0.8256
9	0.8497	0.8446
10	0.8728	0.8639

Source: Own editing from the R statistics calculations

than 0.7. For the Cobb-Douglas function, the largest value was 6.26% greater than the smallest; for the translog function, this value was 5.89%. Therefore, the average efficiency score increased as the total asset portfolio increased, but the change per category was not significant. Hypothesis 6 is rejected.

Table 11 shows the change in efficiency scores according to the total revenue growth. On the basis of the data in Table 11, we can conclude that the companies' average efficiency increased with the total revenue growth. The translog function had higher values in only two cases (1. and 7. rank number) compare to the values of the Cobb-Douglas function, and the differences varied between –0.95% (8. rank number) and 1.16% (1. rank number), which is not a statistically significant difference on least 5% significance level. For both models, the first three categories had an efficiency score less than 0.7. For the Cobb-Douglas function, the largest value was 54.51% greater than the smallest; for the translog function, this value was 51.18%. Thus, the average efficiency score increased as the total revenue increased, and the change per category was significant. Hypothesis 7 is accepted.

CONCLUSION

On the basis of the database used, companies classified according to different criteria (for example, by subsector) had a very heterogeneous picture. Therefore, various additional groupings may be required to reduce heterogeneity. The frontier package of the R statistical system allowed the management of panel models, which enables the combined use of cross-sectional and time series data, which makes it possible to manage the years in the models.

We used analyses of the Cobb-Douglas and translog models. Both models included exogenous variables (*z* variables) and endogenous variables to reduce heterogeneity. According to the results, the translog model had higher efficiency scores than did the Cobb-Douglas model, but the differences were not remarkable. The analysis results indicated that the exogenous variables affected the development of the efficiency scores. The most significant difference occurred in the ranking based on total sales, where the difference between the smallest and largest values exceeded 50%. For a ranking based on total sales revenue, as sales revenue increased, so did enterprise efficiency. The classification according to the workforce also showed that the average efficiency score increased as the number of employees in the company increased.

These findings ground the assumption that considering exogenous variables can improve the analysis results in a heterogeneous database. The use of the Cobb-Douglas and translog models revealed that the translog model provided somewhat better results than did the Cobb-Douglas model. The use of the translog model requires meeting fewer conditions than does the use of the Cobb-Douglas model. At the same time, the regression results were a little more challenging to interpret in the case of the translog model, which does not cause problems if the focus is on evaluating the companies' efficiencies.

On the basis of the research results, we can conclude that the average performance of the examined sector was acceptable, which does not mean that companies do not need to increase performance. For the performance of the Hungarian national economy to improve, all sectors must improve their performance. The research results draw attention to the fact that, on average, there was no significant difference between the individual subsectors in the examined period, but smaller differences can be established. By analysing the efficiency scores, we can also conclude that larger companies can operate with higher efficiency than can smaller ones. This difference also occurred in the workforce classification according to the number of employees, but there was no significant difference between the lowest category and the highest category (Cobb-Douglas function, 6.71%; translog function, 5.21%). The difference was much greater in the classification according to sales, where, compared with the bottom decile, the top decile achieved an efficiency score more than 50% higher. These results draw attention to the fact that smaller companies must im-

https://doi.org/10.17221/125/2022-AGRICECON

prove their efficiency because doing so will also increase the sector's efficiency. If a sector performs better, it increases the national economy's performance.

The research we performed contributed to previous research results in two aspects. On the one hand, most previous research was based on the FADN database, whereas this research was based on the companies' financial statements submitted to the taxation office. On the other hand, we considered several grouping aspects as the *z* variable, which is less common in the literature. These grouping criteria allow better consideration of technological differences.

The limitations of the analysis include the fact that we analysed only a few years, which reduces the validity of the conclusions that can be drawn. Another limitation of the analysis is that additional exogenous variables could probably be included in the model. Perhaps the range of endogenous variables could also be increased. Using value added as a dependent variable in the logarithmic model is a problem because some companies will be excluded. The analysis could be performed with other dependent variables as well. Further statistical testing of the database could also be necessary.

REFERENCES

Abdul I., Wulan Sari D., Haryanto T., Win T. (2022): Analysis of factors affecting the technical inefficiency on Indonesian palm oil plantation. Scientific Reports, 12: 3381.

Abate T.M., Dessie A.B., Adane B.T., Tesfa T., Getu S. (2022): Analysis of resource use efficiency for white cumin production among smallholder farmers empirical evidence from Northwestern Ethiopia: A stochastic frontier approach. Letters in Spatial and Resource Sciences, 15: 213–235.

Aigner D.J., Chu S.F. (1968): On estimating the industry production function. The American Economic Review, 58: 826–839.

Aigner D., Lovell C.A.K., Schmidt P. (1977): Formulation and estimation of stochastic frontier production function models. Journal of Econometrics, 6: 21–37.

Bakucs L.Z., Latruffe L., Fertő I., Fogarasi J. (2010): The impact of EU accession on farms' technical efficiency in Hungary. Post-Communist Economies, 22: 165–175.

Bakucs Z., Fertő I., Vígh E. (2020): Crop productivity and climatic conditions: Evidence from Hungary. Agriculture, 10: 421–432.

Baráth L., Fertő I. (2015): Heterogeneous technology, scale of land use and technical efficiency: The case of Hungarian crop farms. Land Use Policy, 42: 141–150.

Baráth L., Fertő I., Hockmann H. (2020): Technological differences, theoretical consistency, and technical efficiency: The case of Hungarian crop-producing farms. Sustainability, 12: 1147.

Baráth L., Fertő I., Staniszewski J. (2021): Technological heterogeneity in pig farming: A metafrontier approach – Perspectives from Hungary and Poland. Agriculture (Basel), 11: 961–974.

Battese G.E., Coelli T.J. (1992): Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis, 3: 153–169.

Battese G.E., Coelli T.J. (1995): A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics, 20: 325–332.

Bojnec Š., Latruffe L. (2013): Farm size, agricultural subsidies and farm performance in Slovenia. Land Use Policy, 32: 207–217.

Caves R.E., Barton D.R. (1991): Efficiency in US Manufacturing Industries. Cambridge, US, MIT Press: 204.

Coelli T.J., Prasada D.S.P., O'Donell C.J., Battese G.E. (2005): An introduction to efficiency and productivity analysis. New York, US, Springer: 341.

Čechura L., Žáková Kroupová Z., Lekešová M. (2022): Productivity and efficiency in Czech agriculture: Does farm size matter? Agricultural Economics – Czech, 68: 1–10.

Diaz M.A., Sánchez R. (2008): Firm size and productivity in Spain: a stochastic frontier analysis. Small Business Economics, 30: 315–323.

Farrell M.J. (1957): The measurement of productive efficiency. Journal of the Royal Statistical Society (A, general), 120: 253–281.

Hall R.E., Jones C.I. (1999): Why do some countries produce so much more output per worker than others? The Quarterly Journal of Economics, 114: 83–116.

Henningsen A., Coelli T. (2020): Package 'frontier'. Stochastic frontier analysis. Available at http://frontier.r-forge.r-project.org/ (accessed Sept 8, 2021)

Horváthová J., Mokrišová M., Suhányiová A. (2014): Analysis of cost of equity models in calculating economic value added of Slovak businesses. In: International Multidisciplinary Scientific Conferences on Social Sciences and Arts SGEM 2014, Albena, Bulgaria, Sept 2–7, 2014: 35–42.

Giannakis E., Bruggeman A. (2015): The highly variable economic performance of European agriculture. Land Use Policy, 45: 26–35.

Gorton M., Davidova S. (2004): Farm productivity and efficiency in the CEE applicant countries: A synthesis of results. Agricultural Economics, 30: 1–16.

Kim M., Maksimovic V. (1990): Debt and input misallocation. The Journal of Finance, 45: 795–816.

Kalirajan K.P., Shand R.T. (1999): Frontier production functions and technical efficiency measures. Journal of Economic Surveys, 13: 149–172.

Keszthelyi S. (2021): The Results of the Test Farm Information System in 2020 (A tesztüzemi információs rendszer eredmé-

https://doi.org/10.17221/125/2022-AGRICECON

nyei 2020). Budapest, Hungary, Agrárközgazdasági Intézet (Institute of Agricultural Economics): 193. (in Hungarian)

Kovacs K., Szucs I. (2020): Exploring efficiency reserves in Hungarian milk production. Studies in Agricultural Economics, 122: 37–43.

Latruffe L., Fogarasi J., Desjeux Y. (2012): Efficiency, productivity and technology comparison for farms in Central and Western Europe: The case of field crop and dairy farming in Hungary and France. Economic Systems, 36: 264–278.

Lichtenberg F.R., Siegel D. (1990): The effects of leveraged buyouts on productivity and related aspects of firm behavior. Journal of Financial Economics, 27: 165–194.

Meeusen W., van den Broeck J. (1977): Efficiency estimation from Cobb-Douglas production functions with composed error. International Economic Review, 18: 435–444.

Mitsopoulos I., Tsiouni M., Pavloudi A., Gourdouvelis D., Aggelopoulos S. (2021): Improving the technical efficiency and productivity of dairy farms in Greece. Studies in Agricultural Economics, 123: 95–100.

Náglová Z., Šimpachová Pechrová M. (2019): Subsidies and technical efficiency of Czech food processing industry. Agricultural Economics – Czech, 65: 151–159.

Novotná M., Volek T. (2014): Labour as a factor of production in the context of gross value added growth in sector A. Scientia Agriculturae Bohemica, 45: 129–135.

Orea L., Kumbhakar S.C. (2004): Efficiency measurement using latent class stochastic frontier model. Empirical Economics, 29: 169–183.

Palia D., Lichtenberg F. (1999): Managerial ownership and firm performance: A re-examination using productivity measurement. Journal of Corporate Finance, 5: 323–339.

Petrie A.G. (2016): Introduction to Regression and Modeling with R. San Diego, US, Cognella, Inc.: 328.

Pitt M.M., Lee L.F. (1981): The measurement and sources of technical inefficiency in the Indonesian weaving industry. Journal of Development Economics, 9: 43–64.

Richmond J. (1974): Estimating the efficiency of production. International Economic Review, 15: 515–521.

Schmidt P. (1985): Frontier production functions. Econometric Reviews, 4: 289–328.

Sheather S.J. (2009): A Modern Approach to Regression with R. New York, US, Springer Texts in Statistics, Springer Science + Business Media, LLC.: 393.

Simon F., Novák J. (2002): The evaluation of economic situation and comparison of Czech and French agricultural enterprises. Agricultural Economics – Czech, 48: 389–394.

Thomassen M.A., Dolman M.A., van Calker K.J., de Boer I.J.M. (2009): Relating lifecycle assessment indicators to gross value added for Dutch dairy farms. Ecological Economics, 68: 2278–2284.

Van Passel S., Nevens F., Mathijs E., van Huylenbroeck G. (2007): Measuring farm sustainability and explaining differences. Ecological Economics, 62: 149–161.

Zsarnóczai J.S., Zéman Z. (2019): Output value and productivity of agricultural industry in Central-East Europe. Agriculture Economics – Czech, 65: 185–193.

Received: April 24, 2022
Accepted: October 12, 2022