DETERMINANTS OF FARMS' TECHNICAL (IN)EFFICIENCY

Marie Šimpachová Pechrová – Ondřej Šimpach

Abstract

The aim is to find the determinants that influence the technical (in)efficiency of agricultural holdings using different approaches. After the estimation of the production function by True Fixed Effect Model with truncated normal distribution of inefficiency term, the technical efficiency and inefficiency are calculated. First, the explanatory variables are included into the equation of technical inefficiency mean. Second, after the estimation of the model, the (in)efficiency it is included into Tobit regression as explained variable.

There were 1708 observations for 517 firms for 2013–2016. Average technical efficiency was 85.7% and inefficiency 47.4%. When the consumption of material and energy increases, the mean of inefficiency decreases. According to the Tobit model, the efficiency increases, and the inefficiency decreases. The higher are the fixed assets, the higher is the mean of inefficiency or the lower is the efficiency and higher inefficiency. The number of employees influence the efficiency negatively. Coefficient for land is statistically significant only in Tobit model with inefficiency. If the acreage increases, the inefficiency also increases.

The best solution for our data is to calculate the technical inefficiency and include explanatory variables into Tobit model where all determinants are statistically significant. Nevertheless, their influence is very mild.

Key words: farm, stochastic frontier analysis, technical efficiency

JEL Code: C23, C24

Introduction

Technical efficiency concept was defined by Farrell (1957) as the ability of firm to produce as large as possible output from given set of inputs. Other way round it can be defined as the given output that can be reached with minimum inputs. It represents a technical transformation of inputs to outputs of a firm. All firms are compared among each other according to the efficiency of this transformation process. Technical efficiency takes values between 0 and 1 (or 0% to

100%). "The company that is 100% efficient lies on the efficiency frontier. "If TE < 1, it indicates that the output of the enterprise lies below the production frontier and is in a state of technical inefficiency." (Sun et al., 2017)

There are two approaches how the technical efficiency can be calculated. Non-parametrical approach "DEA assumes that there are no random fluctuations from the efficient frontier, i.e. all deviations are considered inefficiency." (Chen, 2007). Parametric approach estimates the parameters of the production function, so its type of function must be pre-defined. One of the methods is Stochastic Frontier Analysis (SFA). A main advantage is that it isolates the influence of factors other than inefficient behaviour, thus correcting the possible upward bias of inefficiency from the deterministic methods. (Chen, 2007). Mostly used functions are Cobb-Douglas or translog. SFA can be applied on panel data, we can consider time-invariant or time variant models.

There are two approaches how to assess the determinants of the technical efficiency. First, the determinants can be included into the function of mean of technical inefficiency or to the function of variance of the technical efficiency. In this case, the inefficiency term has to have truncated normal distribution. There are functions of mean and variance of technical inefficiency where explanatory variables can be included. Truncated normal distribution for the inefficiency term was proposed by Pitt and Lee (1981) and one-step procedure of estimation of the parameters of the inefficiency model by Battese and Coelli, (1995). For example, Song and Chen (2019) used truncated normal distribution and included explanatory variables into the inefficiency model that was estimated at the same time as translog production function. Pechrová and Vlašicová (2013) tried to assess the impact of subsidies and whether the farm is organic or biodynamic on the technical efficiency, so they included the explanatory variables (constant and various types of subsidies) to the inefficiency variance function. Liu et al. (2019) included a vector of the environmental factors into the mean of the pre-truncated normal variable in order to investigate the influence of those environmental variables on inefficiency.

Second, the efficiency can be an explanatory variable in Tobit regression that is specialized on censored samples (TE takes values from 0 to 1). For example, Scippacercola and Ambra (2014) adopted the SFA to estimate the efficiency and a Tobit regression to discuss which factors might affect it. "In this approach the efficiency scores are first estimated using DEA and in a second stage either OLS or Tobit is used in regressions of these efficiency estimates on a number of contextual variables." (Banker, Natarajan and Zhang, 2019). For example, Lerida-Navarro, Nombelah and Tranchez_Martin (2019) uses the efficiency scores obtained from the DEA as explained variable in Tobit model and introduced various factors (country characteristics, characteristics of the railways' network and the progress in rail reforms) to assess the determinants of the technical efficiency of European railways. Similarly, two-stage efficiency analysis based on DEA efficiency and the Tobit model was used by Fitzova and Matulova (in press) to identify conditions important for efficient urban public transport.

The aim of the paper is to find the determinants that influence the technical (in)efficiency of agricultural holdings using those two above stated approaches.

1 Data and Methods

The explained variable (y_t of particular farm i in time t) in the production function were sales of own products and services (in thous. CZK) that were adjusted by agricultural producers' price index (2010 = 100) as the inflation changes the value of the production in time. Production factor material ($x_{1,it}$) was represented by the amount of consumed material and energy by i^{th} farm in time t. Capital ($x_{2,it}$) consisted of fixed assets of i^{th} farm in time t. Both variables were deflated by the industrial producers' price index (2010 = 100). Labour ($x_{3,it}$) was represented by number of employees. These data were taken from Albertina database. The acreage in hectares (input land – $x_{4,it}$) was obtained from LPIS database. There were 1708 observations for 517 farms (3.3 per farm) from years 2013 to 2016. The description of the variables is in Tab. 1.

Tab. 1: Description of the variables

Variable	Mean	Std. dev.	Median	Min	Max
y_{it} – sales of own products and services (adjusted) [1000 CZK]	58453	64245	40303	27	498372
x_{1it} – material and energy consumption (adjusted) [1000 CZK]	29057	30932	21404	1	315756
x_{2it} - fixed assets (adjusted) [1000 CZK]	69011	90522	56698	1	1200000
x_{3it} – number of employees [-]	42	42	38	3	225
x_{4it} – acreage [ha]	49447	1156	739	0	10381

Source: own elaboration

We considered Cobb-Douglas production function (1) because of ease of estimation and possibility to interpret the estimated coefficients as elasticities (power function can be linearized by natural logarithms).

$$y_{it} = x_{1,it}^{\beta_1}.x_{2,it}^{\beta_2}.x_{3,it}^{\beta_3}.x_{4,it}^{\beta_4}.e^{v_{it}-u_{it}} \qquad \ln y_{it} = \sum_{k=1}^{K} \beta_k \ln x_{k,it} + \varepsilon_{it} \qquad (1)$$

where y_{it} is the output (production) of farm i (i = 1, 2, ..., N, where N is total number of farms) in time t (t = 1, 2, ..., T, where T stays for a time). $x_{k,it}$ (k = 1, 2, ..., K, where K is total number of production factors) stands for the input k of firm i in time t; β_k are the estimated parameters of inputs; ε_{it} is the synthetic error term ($\varepsilon_{it} = v_{it} - u_{it}$, where v_{it} marks the random systematic error). This was introduced by Aigner, Lovell and Schmidt (1977) who proposed a method that distinguishes productive inefficiency from other sources of disturbance. The second term, u_{it}, is a non-negative random variable, which is assumed to account for the existence of technical inefficiency of production of the i^{th} firm at the t^{th} period of observation. It stands for the error caused by the technical inefficiency and can be interpreted as the percentage deviation of observed performance (y_{it}) from the firm's own frontier (Greene, 2005). It measures the distance from the frontier. (Kumbhakar and Wang, 2005). In our model, the values of v_{it} and u_{it} change over time and across different farms in the model.

The distribution of the inefficiency term was assumed to be truncated normal ($u_{it} \sim N^{+}(\mu, \sigma_u^2)$) – truncated at mean μ with variance σ_u^2 and stochastic noise normal distribution that is independently and identically distributed ($v_{it} \sim N(0, \sigma_v^2)$) with 0 mean and variance σ_v^2.

We follow the approach of Battese and Coelli (1995): "The inefficiency effects are assumed to be independently distributed as truncations of normal distributions with constant variance, but with means which are a linear function of observable variables". We considered the inefficiency term to have constant variance, i.e. being homoscedastic. On the other hand, the mean of technical inefficiency contains explanatory variables to capture the heterogeneity among firms. This is one way, how to assess the determinants of the technical inefficiency. The inefficiency effects are assumed to be defined by (2).

$$u_{it} = \delta_0 + \delta_1 x_{1,it} + \delta_2 x_{2,it} + \delta_3 x_{3,it} + \delta_4 x_{4,it} \qquad (2)$$

where δ_0 is constant and δ_k (where k = 1, ... K) are coefficients of explanatory variables $x_{k,it}$. Those variables are the same as in the production function. These factors that cause inefficiency can be improved by firm's effort. However, heterogeneity, which does not vary with time, is beyond the control of enterprises. From the perspective of model analysis, the reason for the existence of heterogeneity lies in the following aspects: the omission of time-invariant input variables, the ignorance of explanatory variables that are difficult to quantify, and no data available for the explanatory variables (Lin et al., 2010).

A production function was put to True fixed-effects model elaborated by Greene (2005). It is one of the time-varying models – i. e. firms' technical inefficiency can improve over time. The parameters of stochastic frontier function are estimated by the maximum likelihood method. After that, the inefficiency and efficiency were estimated using Jondrow et al. (1982)

method. So-called JLMS method measures the contribution of u_{it} to ε_{it}. Their estimator of u_{it} is calculated as (3).

$$E\left[u_{it}\left|\varepsilon_{it}\right.\right]=\frac{\sigma\lambda}{1+\lambda^{2}}\left[\frac{\phi(a_{it})}{1-\Phi(a_{it})}-a_{it}\right], \qquad (3)$$

where $\varepsilon_{it}=v_{it}-u_{it}$, $\sigma=\sqrt{\sigma_v^2+\sigma_u^2}$ is standard deviation that is sum of variance of stochastic and inefficiency term, $\lambda=\frac{\sigma_u}{\sigma_v}$, $a_{it}=\pm\frac{\varepsilon_{it}\lambda}{\sigma}$, $\phi(a_{it})$ denote the distribution function in a_{it} of standardized normal distribution and $\Phi(a_{it})$ cumulative distribution function in a_{it}. Technical efficiency is then calculated as exp[-$E(u_{it}|\varepsilon_{it})$].

Second, there is a two-stage approach. After the estimation of the technical inefficiency (TI) and efficiency (TE) both are included into Tobit regression (for censored explanatory variable) as explained variables and the determinants are explanatory variables (4). Censoring limits may be fixed for all observations or vary across observations. We chose the limits from 0 to 1 as the technical efficiency and inefficiency take values from this interval.

$$TI/TE=\alpha_0+\sum_{k=1}^{K}\alpha_k x_{k,it}+\varepsilon', \qquad (4)$$

where α_0 is a constant, α_k are the coefficients of explanatory variables k (k = 1, 2, ..., K, where K is total number of production factors) and ε' is a stochastic term (different from that in (2)). Both procedures are done and compared. Calculations were done in Stata/IC 15.1.

2 Results and discussion

The estimated model as a whole was statistically significant, because p-value = 0.00 for Wald χ^2 test was lower than level of significance α = 0.05. Wald $\chi^{2\,[4]}$ = 1803.56 and of log likelihood was equal to 315.71. Estimated parameters are presented in Tab. 2. Increase of consumed material and energy brings increase of production by 0.29%, increase of fixed assets by 0.26%, increase of number of employees by 0.55% and of land by 0.12%. The highest influence on production had the number of employees.

In the mean of the inefficiency function, all coefficient with exception that one for land are statistically significant at α = 0.05 level. Higher consumption of material and energy cause decrease of technical inefficiency by a very small number. Increase of fixed assets and number of employees cause the increase of the technical inefficiency.

Average technical inefficiency 47.35% was and efficiency 85.69%. Median of inefficiency was lower than mean (37.92%) and of efficiency higher (89.99%). Values of skewness of inefficiency (1.7776) and efficiency (-1.6308) and kurtosis (9.4353; 6.0365, resp.) show that the variables are not from normal distribution.

Tab. 2: True fixed-effects model with truncated normal distribution of inefficiency term

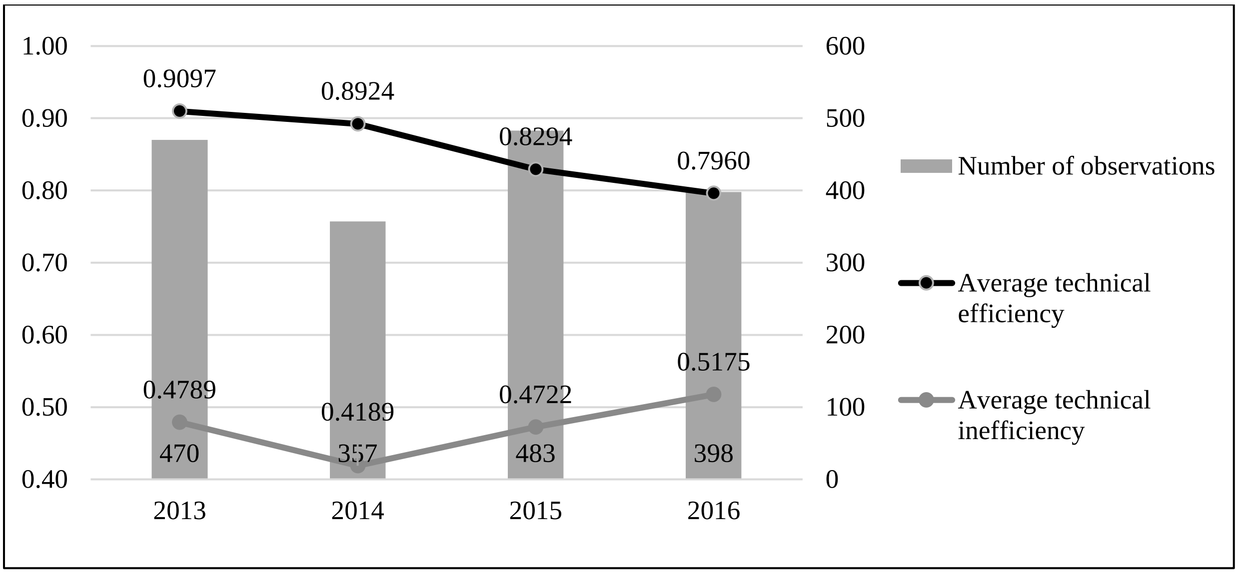
	Coeff. (Std. error) $^{p\text{-value}}$	[95% confidence interval]	
Frontier			
β_1 [ln x_1]	0.2982 (0.0222)***	0.2547	0.3416
β_2 [ln x_2]	0.2625 (0.0223)***	0.2187	0.3063
β_3 [ln x_3]	0.5518 (0.0398)***	0.4738	0.6298
β_4 [ln x_4]	0.1182 (0.0261)***	0.0671	0.1694
Variance of statistical noise function			
constant	-3.2089 (0.0410)***	-3.2893	-3.1284
σ_u	0.0053 (0.0492)***	0.0000	430194
σ_v	0.2010 (0.0041)***	0.1931	0.2093
λ	0.0263 (0.0506)***	-0.0729	0.1255

	Coeff. (Std. error) $^{p\text{-value}}$	[95% confidence interval]	
Mean of inefficiency term			
δ_0	0.1004 (0.0752)	-0.0470	0.2477
δ_1 [x_1]	-5.50 x 10^{-6} (1.12 x 10^{-6})***	0.0000	0.0000
δ_2 [x_2]	2.39 x 10^{-6} (4.44 x 10^{-7})***	0.0000	0.0000
δ_3 [x_3]	0.0054 (0.0008)***	0.0038	0.0071
δ_4 [x_4]	0.0001 (0.0007)	-0.0012	0.0014
Variance of inefficiency term			
constant	-10.4840 (18.5860)	-46.9119	25.9440

Source: own elaboration; Note: *** $\alpha = 0.01$

The highest efficiency was in 2013, but the inefficiency was also very high in that year. The lowest inefficiency was noted in 2014 and the efficiency was also relatively high. A year later, the inefficiency increased and efficiency decreased.

Fig. 1. Development of technical inefficiency and efficiency



Source: own elaboration

Tobit models was statistically significant, because LR $\chi^{2[4]}$ = 12446.26 in case of inefficiency and LR $\chi^{2[4]}$ = 57.38 in case of efficiency had both p-value equal to 0, that was lower than the level of significance $\alpha = 0.05$ (see Tab. 3). However, all coefficients were statistically significant only in inefficiency model. Coefficients for the number of employees and land were statistically insignificant in efficiency model. When the consumption of material and energy increases, the mean of technical inefficiency decreases by very small number. According to the Tobit model, the technical efficiency increases by also very small number (1.43×10^{-6}) and the technical inefficiency decreases by 5.38×10^{-6}.

The higher are the fixed assets, the higher is the mean of technical inefficiency or the lower is the technical efficiency and higher technical inefficiency. In all cases, it is very small number. The number of employees influence the technical efficiency negatively - the higher is the number, the higher is the technical inefficiency, lower technical efficiency and higher technical inefficiency, but in case of the Tobit model for technical efficiency, the coefficient is not statistically significant. Coefficient for land is not statistically significant in the main and Tobit model for efficiency, so we cannot conclude whether the higher acreage (size of the farm) means that the farm is more or less technically efficient based on those models. Only in Tobit model with explained inefficiency all coefficients were statistically significant. For the land, if the acreage increases, the technical inefficiency increases by 9.75×10^{-5}.

Tab. 3: Tobit regression with explained inefficiency and efficiency

inefficiency	**Coeff. (Std. error)** $^{\text{p-value}}$	**[95% confidence interval]**	
α_0	0.1016 (0.0003)***	0.1011	0.1021
x_1	-5.38×10^{-6} (1.22×10^{-8})***	-5.41×10^{-6}	-5.36×10^{-6}
x_2	2.37×10^{-6} (4.27×10^{-9})***	2.36×10^{-6}	2.38×10^{-6}
x_3	0.0054 (8.72×10^{-6})***	0.0054	0.0054
x_4	0.0001 (2.46×10^{-7})***	0.0001	0.0001
Var(e.ineff)	4.13×10^{-5}	3.85×10^{-5}	4.43×10^{-5}

efficiency	**Coeff. (Std. error)** $^{\text{p-value}}$	**[95% confidence interval]**	
α_0	0.8352 (0.0059)***	0.8237	0.8467
x_1	1.43×10^{-6} (2.77×10^{-7})***	8.87×10^{-7}	1.98×10^{-6}
x_2	-2.76×10^{-7} (7.84×10^{-8})***	-4.30×10^{-7}	-1.22×10^{-7}
x_3	-0.0001 (0.0002)	-0.0004	0.0002
x_4	7.21×10^{-6} (5.36×10^{-6})	-3.30×10^{-6}	1.77×10^{-5}
Var(e.eff)	0.0269 (0.0009)***	0.0252	0.0288

Source: own elaboration

Results across all models are consistent. Increase of material an energy consumption cause decrease of mean of technical inefficiency and inefficiency; and also increase of efficiency. Increase of fixed assets cause increase of mean of inefficiency and inefficiency and decrease of efficiency. Increase of number of employees increase of mean of inefficiency and inefficiency and decrease of efficiency. If the acreage increases, the technical inefficiency increases. All estimated coefficients have expected results, that the increase of the usage of production factors cause increase of inefficiency. Only exception was consumption of material where surprisingly the efficiency increased when was used more of it. However, by a very small number.

We can conclude that Tobit model with explained inefficiency had the best results as all coefficients were statistically significant. Similar approach as ours was adopted e. g. by Scippacercola and Ambra (2014) who considered two-stage procedure as suitable for the estimation of efficiency of secondary schools. However, the influence of the determinants was only mild, so other determinants of technical (in) efficiency must be examined in the future.

Conclusion

The aim of the paper was to find the determinants that influence the technical (in)efficiency of agricultural holdings using different approaches. After the estimation of the production function by True Fixed Effect Model with truncated normal distribution of inefficiency term, the technical efficiency and inefficiency are calculated. Their determinants were assessed: (1) the explanatory variables were included into the equation of mean of technical inefficiency, (2) after the estimation of the model, the (in)efficiency it is included into Tobit regression (for censored samples) as explained variable and the determinants as explanatory variables.

Average technical efficiency was 85.69% and inefficiency 47.35%. The highest influence on production had the number of employees. When the consumption of material and energy increased, the mean of technical inefficiency decreased by very small number. According to the Tobit model, the technical efficiency also increased, and the technical inefficiency decreased. The higher were the fixed assets, the higher was the mean of technical inefficiency or the lower was the efficiency and higher inefficiency. The number of employees influenced the technical efficiency negatively – the higher was the number, the higher was the mean inefficiency, lower efficiency and higher inefficiency. Coefficient for land was statistically significant only in model with explained inefficiency. If the acreage increased, the inefficiency also increased.

For our data, it is the best solution to calculate the technical inefficiency and include its explanatory variables into the Tobit model. In this case, all determinants are statistically significant. Nevertheless, their influence is very mild. Hence, in future research, other determinants that influence the technical (in)efficiency shall be examined.

Acknowledgment

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

References

Aigner, D., C.A., Lovell, K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6*, 21–37. doi: 10.1016/0304-4076(77)90052-5

Banker, R., Natarajan, R., Zhang, D. (2019). Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using Data Envelopment Analysis: Second stage OLS versus bootstrap approaches. *European Journal of Operational Research, 278*, 368–384. doi: 10.1016/j.ejor.2018.10.050

Battese, G. E. & Tessema, G. A. (1993). Estimation of stochastic frontier production functions with time-varying parameters and technical efficiencies using panel data from Indian villages. *Agricultural Economics, 9*, 313–333. doi: 10.1016/0169-5150(93)90020-D

Farrell, M. J. (1957). The measurement of Productive efficiency. *Journal of the Royal Statistical Society, 120*, part III. Series A (General), 253–290. doi: 10.2307/2343100

Fitzová, H. &a Matulová, M. (*in press*). Comparison of urban public transport systems in the Czech Republic and Slovakia: Factors underpinning efficiency. *Research in Transportation Economics.* doi: 10.1016/j.retrec.2020.100824

Chen, Ch. (2007). Applying the stochastic frontier approach to measure hotel managerial efficiency in Taiwan. *Tourism Management, 28,* 696–702 doi: 10.1016/j.tourman.2006.04.023

Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*(2-3), 233–238. doi: 10.1016/0304-4076(82)90004-5

Kumbhakar, S. C., & Wang, H.-J. (2005). Estimation of growth convergence using a stochastic production frontier approach. *Economics Letters, 88,* 300–305 doi: 10.1016/j.econlet.2005.01.023

Lerida-Navarroa, C., Nombelab, G., & Tranchez-Martin, J. M. (2019). European railways: Liberalization and productive efficiency. *Transport Policy*, *83*, 57–67. doi: 10.1016/j.tranpol.2019.09.002

Liu, X., Pollitt, M. G. Xie, B., & Liu, L. (2019). Does environmental heterogeneity affect the productive efficiency of grid utilities in China? *Energy Economics*, *83*, 333–344. doi: 10.1016/j.eneco.2019.07.001

Pechrová, M., &Vlašicová, E. (2013). Technical Efficiency of Organic and Biodynamic Farms in the Czech Republic. *Agris on-line Papers in Economics and Informatics*, *V*.(4), 143–152.

Pitt, M., & Lee, L. (1981). The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics*, *9*, 43–64. doi: 10.1016/0304-3878(81)90004-3

Scippacercola, S., & D’Ambra, L. (2014). Estimating the Relative Efficiency of Secondary Schools by Stochastic Frontier Analysis. *Procedia Economics and Finance*, *17*, 79–88. doi: 10.1016/s2212-5671(14)00881-8

Sun, C. et al. (2017). A comparative study on the production efficiencies of China's oil companies: A true fixed effect model considering the unobserved heterogeneity. *Journal of Cleaner Production*, *154*, 341–352. doi: 10.1016/j.jclepro.2017.03.222

Contact

Ing. Marie Šimpachová Pechrová, Ph.D.
Institute of Agricultural Economics and Information
Mánesova 1453/75, 120 00 Prague 2, Czech Republic
simpachova.marie@uzei.cz

Ing. Ondřej Šimpach, Ph.D.
University of Economics Prague
nám. W. Churchilla 1938/4, 130 67 Prague 3, Czech Republic
ondrej.simpach@vse.cz