https://doi.org/10.17221/347/2020-AGRICECON

Measurement of technical efficiency in the case of heterogeneity of technologies used between firms – Based on evidence from Polish crop farms

*Jerzy Marzec, Andrzej Pisulewski**

Department of Econometrics and Operational Research, Institute of Quantitative Methods in Social Sciences, Cracow University of Economics, Kraków, Poland

**Corresponding author: andrzej.pisulewski@uek.krakow.pl*

Citation: Marzec J., Pisulewski A. (2021): Measurement of technical efficiency in the case of heterogeneity of technologies used between firms – Based on evidence from Polish crop farms. Agri. Econ. – Czech., 67: 152–161.

Abstract: In the present study, we have investigated several competing stochastic frontier models which differ in terms of the form of the production function (Cobb-Douglas or translog), inefficiency distribution (half-normal or exponential distribution) and type of prior distribution for the parameters (hierarchical or non-hierarchical from the Bayesian point of view). This last distinction corresponds to a difference between random coefficients and fixed coefficients models. Consequently, this study aims to examine to what extent inferences about estimates of farms' efficiency depend on the above assumptions. Moreover, the study intends to investigate how far the production function's characteristics are affected by the choice of the type of prior distribution for the parameters. First of all, it was found that the form of the production function does not impact the efficiency scores. Secondly, we found that measures of technical efficiency are sensitive to distributional assumptions about the inefficiency term. Finally, we have revealed that estimates of technical efficiency are reasonably robust to the prior information about the parameters of crop farms' production technology. There is also a resemblance in the elasticity of output with respect to inputs between the models considered in this paper. Additionally, the measurement of returns to scale is not sensitive to model specification.

Keywords: Bayesian approach; crop farms; random coefficients; stochastic frontier analysis

In empirical studies of economic efficiency, where a frontier production function model with panel data is employed, a restrictive assumption is made that all firms, or decision making units (DMUs), must share precisely the same technological possibilities. However, DMUs do not themselves form a homogeneous sample and vary greatly in their technological capabilities. Therefore, modelling heterogeneity is important, especially in the case of agricultural data, because farms usually conduct several agricultural activities, which depend on various factors, both specific and general (external). As pointed out by Tsionas (2002), in such a case assuming that firms share the same technology will result in the misleading measurement of efficiency. Therefore, when a researcher decides to focus on a specific type of agricultural production, he has to choose criteria to distinguish it and run an analysis on subsamples of homogenous technology. However, this choice will always be arbitrary. Therefore, a non-arbitrary approach to handling sample heterogeneity should be used instead. In the literature, there are several such approaches to doing this. However, there are two of them which are the most common. The first one is the random coefficients stochastic frontier (SF) model (Tsionas 2002; Greene 2005). The special case of the aforementioned model is the true random effects (TRE) stochastic frontier model. In this case, only the constant term is random; thus, Greene (2005)

Supported by the subsidy granted to the Cracow University of Economics.

https://doi.org/10.17221/347/2020-AGRICECON

originally called this model a random constant model. An alternative approach is a latent class stochastic frontier model (Caudill 2003; Orea and Kumbhakar 2004; Greene 2005). As noted by Greene (2005) and Emvalomatis (2012), in a random coefficients setting, firms' heterogeneity is captured by continuous variation in technical parameters, while in the latent-class model, this continuous variation is achieved by a discrete approximation. Therefore, in the present study, we employ the more general approach using the so-called random coefficients stochastic frontier model and its special case, i.e. the TRE model. The results from these models will be subsequently contrasted with the fixed coefficients stochastic frontier model.

The stochastic frontier models, and especially the random coefficients stochastic frontier models, have a hierarchical structure, in which case, a Bayesian approach is a particularly convenient tool since it allows exact estimation of parameters, model selection and description of uncertainty. The Bayesian approach to stochastic frontier models was first proposed by van den Broeck et al. (1994) and Koop et al. (1997). Since then, it has been successfully applied in a number of studies and further developed (Areal et al. 2012; Marzec and Pisulewski 2017; Skevas et al. 2018).

A relatively small number of applications of the random coefficients stochastic frontier model in farm efficiency analysis are by no means a new concept. Among the existing analyses, it is mainly German dairy farms that have been studied, e.g. Emvalomatis (2012), Skevas (2019). Models with random coefficients have also been used in a study of Czech crop farms by Čechura (2010) and an analysis of U.S. agricultural state-level data by Njuki et al. (2019). Moreover, Čechura et al. (2017) and Baráth et al. (2018) have used the special case of random coefficients models, i.e. the Alvarez et al. (2004) model, in the analysis of European dairy farms and Slovenian farms, respectively. The above-mentioned studies usually provide a comparison between fixed and random coefficients models. However, this distinction does not necessarily have to be the crucial one. We argue that the other assumptions made in stochastic frontier models, such as inefficiency distribution and the form of production function, can also affect efficiency measurement and production characteristics. Therefore, this paper aims to examine the robustness of the results obtained with respect to three aspects: an approach to heterogeneity (hierarchical and non-hierarchical prior distribution), distribution of the inefficiency term and the form of the production function.

Consequently, on the one hand, our study contributes to the literature on modelling technical efficiency (TE) of heterogeneous data. On the other hand, it provides a measurement of productive efficiency in Poland's particular agricultural sector with models that account for heterogeneity, contributing to the empirical literature on this subject. Moreover, the chosen time-span (2004–2011) of the sample allows the direction of technical efficiency changes to be compared between Polish crop farms and those in other European Union countries for a similar time period (Čechura et al. 2015).

MATERIAL AND METHODS

Data on Polish crop farms. To measure the productive efficiency of heterogeneous farms, data on 660 crop farms from the Polish Farm Accountancy Data Network (FADN 2020) was employed. These farms were observed over a period of eight years (2004–2011). Subsequently, the variables of the production function model were put together based on this data. The definition of these variables is based on other studies in the field crop sector in which FADN data was used (Latruffe et al. 2004; Bojnec and Latruffe 2009; Zhu and Lansink 2010). Therefore, the output (Q) is specified as the deflated total net farm revenues from sales (deflated with the year 2004 as the base period), excluding the value of feed, seeds and plants produced within the farm. The four inputs of production are physical capital (K), total labour (L; hours), total utilised agricultural area (A; ha) and materials (M). It is worth mentioning that the latter input consists of several subcategories (seeds and plants, fertilisers, crop protection, among others). To get the real value of the variables, i.e. of Q, K and M, price indexes provided by the Central Statistical Office of Poland are used as deflators. The exact definitions of these quantities, as well as information about the deflators, were presented by Marzec and Pisulewski (2017), Pisulewski and Marzec (2019).

Table 1 summarises the descriptive statistics for the variables used in the main study. The arithmetic mean area of land per farm is 43 ha. Crop farms in European countries are generally larger (Pisulewski and Marzec 2019).

Analytical framework. In the present study, we use the random coefficients stochastic frontier model. The general concept of such models dates back to the works of Hildreth and Houck (1968) and Swamy (1970). It is noteworthy that this type of model has found broad application, for example, in marketing (Rossi et al. 2005),

https://doi.org/10.17221/347/2020-AGRICECON

Table 1. Descriptive statistics for data set

Variable*	Mean**	Percentile				
		5th	25th	50th	75th	95th
Output (in thousand EUR)	29	6	15	28	56	143
Capital (in thousand EUR)	56	13	30	56	104	252
Labour (in hours)	4 056	1 826	2 900	3 938	5 214	11 013
Materials (in thousand EUR)	20	5	10	19	36	95
Agricultural area (in ha)	43	10	21	40	83	220

*Figures were first deflated (with base year 2004) and then converted at an exchange rate of PLN 4.15 to EUR 1
**Descriptive statistics for output and input variables were calculated on the logarithmic scale and then transformed back to the original scale
Source: Pisulewski and Marzec (2019); FADN (2020)

and therefore models of this type are referred to under different names, i.e. as hierarchical models, mixed models or multilevel models (Greene 2012). From the Bayesian point of view, the specification of a hierarchical model includes not only the likelihood function but also two stages of prior distributions, i.e. for unit-level parameters and for the common parameters.

The stochastic frontier model with random coefficients was first introduced by Kalirajan and Obwona (1994) and, in the case of i^{th} object, this model takes the following form:

$$y_i = W_i\gamma + X_i\beta_i + v_i - u_i \tag{1}$$

where: y_i – T-elementary vector which includes the natural log of the observed output for firm i ($i = 1, ..., N$); X_i – matrix that contains k explanatory variables which come from T time periods; β_i – (column) vector of k parameters.

Furthermore, β_i has a k-variate normal distribution with expected value β and covariance matrix Ω, which is symmetric and positive-definite. Consequently, the coefficients β_i are treated as hidden random variables. In that case, the coefficients vary between individuals. Moreover, the Equation (1) includes coefficients γ, which correspond to the explanatory variables grouped together in vector W_i, which are not randomly distributed. An example of such variables could be the constant term or the trend. The conventional fixed coefficients SF model is obtained when all explanatory variables are included in W_i. Additionally, u_i and v_i are T-elementary vectors representing inefficiency and the random term, respectively. Component $u_{it} \geq 0$ is referred to as inefficiency, and so the output-oriented technical efficiency (TE) score is calculated as $\exp(-u_{it})$. The conventional assumption is that the error term and the inefficiency term are independently and identically distributed across units and time. In the study by Aigner et al. (1977), the inefficiency term is derived from a normal distribution truncated above at zero, $N^+\left(0,\sigma_u^2\right)$, or it has an exponential distribution with mean λ, hereafter referred to as $\mathrm{EXP}(\lambda)$. Other commonly adopted distributions are the truncated-normal and gamma distributions.

The feature that distinguishes model defined by Equation (1) is the fact that the parameters included in the production function (β) are firm-specific. In the case of the Cobb-Douglas production function, it means that output elasticities with respect to inputs are different for each firm. The coefficients β_i are identically and independently normally distributed, thus have the same expected value and non-singular covariance matrix. Therefore, in practice, farms are characterised by individual parameters that should be similar between the farms. Both parameters, the vector β and the Ω matrix, which define the probability distribution of β_i, are unknown and they must therefore be estimated. It is noteworthy that in the present study, the covariance matrix Ω is not diagonal. Restricting the Ω matrix to be diagonal is a practice quite often used in the empirical literature (Baráth et al. 2018; Njuki et al. 2019). On the one hand, this restriction facilitates the estimation procedure because there are fewer elements in the matrix to be estimated (only k), while, in the case of the full covariance matrix, there are $0.5 \times k \times (k + 1)$ elements. On the other hand, this assumption excludes any interrelated changes in the parameters. If the multi-inputs production function is considered in order to correlate the technical rate of substitution of different pairs of production factors, it is necessary to correlate the elements of the vector β_i.

In the Bayesian approach, the statistical model is defined by joint distribution for the data set and prior distributions for the parameters model. Therefore, there

https://doi.org/10.17221/347/2020-AGRICECON

is a possibility via prior distributions to introduce some information about parameters, which is dictated by the economic theory or specificity of the considered issue. First of all, a normal prior with mean μ_β and covariance matrix V_β is used for β_i. The hyperparameter μ_β is chosen so that the expected value of the output elasticity with respect to each production factor at the geometric mean of each input is 0.25, and consequently, there are constant returns to scale. At the same time, the matrix V_β is assumed to be diagonal with 10 on diagonal. Therefore, a high variance for the elements of β_i is assumed. Subsequently, the Wishart prior is assumed for the precision matrix Ω^{-1} (Emvalomatis 2012). Assuming this type of distribution guarantees the positive-definiteness of Ω^{-1}, as well as of the covariance matrix Ω. The next parameter is the precision of white noise term $\left(\sigma_v^{-2}\right)$. The standard procedure is to assume the gamma distribution for this parameter. The hyperparameters are chosen so that the prior distribution is reasonably diffuse (Koop et al. 1997; Tsionas 2002).

Furthermore, in the present study, we consider two types of distribution for the inefficiency term. In the case of hyperparameters defining the prior distribution for the inefficiency term, described by the exponential distribution or half-normal distribution, it is assumed that the reciprocals of these parameters (λ^{-1} and σ_u^{-2}) have a gamma distribution. The four hyperparameters are chosen so that the median of the prior distribution for efficiency is approximately 0.8 (Koop et al. 1997).

The special case of the aforementioned random coefficients model is the true random effects (TRE) model, which was proposed by Greene (2005). The model takes the following form:

$$y_{it} = w_{it}\gamma + \alpha_i + v_{it} - u_{it} \tag{2}$$

where: $w_{it}\gamma$ – equivalent to the formula given previously in Equation (1); α_i – heterogeneous intercept that represents the effects of hidden variables specific to firm i in the same fashion over time.

However, unlike inefficiency, this random variable can assume both negative and positive values because it is related to the total impact of other unobserved characteristics of firms. Following Feng and Zhang (2012), who presented a Bayesian approach to the TRE model, a normal prior: $\alpha_i \sim N\left(0, \sigma_\alpha^2\right)$ is assumed, while the prior distribution of σ_α^{-2} is gamma, just as for scale hyperparameters discussed above. For other models, a normal prior is used for a common intercept. Almost all of the aforementioned prior distributions reflect very weak prior knowledge or restriction and allow the data to "speak for themselves".

RESULTS AND DISCUSSION

The foundation of efficiency measurement is to identify the corresponding production possibility set and compare it with farms' decisions. The considered models differ in terms of the form of the production function [Cobb-Douglas (CD) or translog], inefficiency distribution (half-normal or exponential distribution) and type of prior distribution for the parameters (hierarchical or non-hierarchical from the Bayesian point of view). This last distinction corresponds to a difference between random coefficients and fixed coefficients models. Consequently, we have considered ten models, the detailed assumptions of which are presented in Table 2.

The parameters of the stochastic frontier models are estimated using the Gibbs sampler algorithm. We did an initial run of 100 000 iterations and discarded the first 50 000 draws as "burn-in". After the "burn-in", these 50 000 draws were used for estimating the marginal posterior distribution of the quantity of interest. We then applied the Bayesian model comparison approach to choose a model which corresponds to the maximum posterior probability. In particular, in the present study, the Chib (1995) method is used to compute the marginal likelihood. Statistical analysis of the data was done using BayES^{TM} software (Bayesian Econometrics Software, version 2.4).

As far as the form of production function is considered, it should be noted, based on posterior odds presented in Table 2, that the translog production function is favoured by the data as long as a non-hierarchical prior is assumed for technology parameters. However, when a hierarchical prior is assumed for technology parameters, then the CD production function is favoured by the data. This result seems to suggest that the CD specification in the hierarchical prior framework achieves more parsimonious parameterisation than the other specifications considered. Subsequently, it can be seen that the exponential distribution of inefficiency seems to be more favoured than a half-normal distribution. Moreover, Table 2 shows that the models with a hierarchical structure (i.e. with parameters varying between farms) are preferred over the ones with a non-hierarchical prior. Finally, it should be noted that the most favoured model based on the data is the true random effects model with the translog production function and a varying intercept (M_9), while the second one is the Cobb-Douglas model (M_5) with firm-specific

https://doi.org/10.17221/347/2020-AGRICECON

Table 2. Choice of model [equal prior model probabilities, i.e. $P(M_r) = 0.1$]

M_r	Model specification	Type of prior for production function parameters	Log-marginal likelihood	Bayesian posterior odds ratios	Model rank
M_1	CD, u_{it} exponential, constant β	non-hierarchical	–1 513.780	≈ 0	9
M_2	CD, u_{it} half-normal, constant β	non-hierarchical	–1 552.560	≈ 0	10
M_3	Translog, u_{it} exponential, constant β	non-hierarchical	–1 462.470	≈ 0	7
M_4	Translog, u_{it} half-normal, constant β	non-hierarchical	–1 510.230	≈ 0	8
M_5	CD, u_{it} exponential, firm-specific β	hierarchical	–745.903	5.04×10^{-13}	2
M_6	CD, u_{it} half-normal, firm-specific β	hierarchical	–793.269	1.35×10^{-33}	4
M_7	Translog, u_{it} exponential, firm-specific β	hierarchical	–819.660	4.68×10^{-45}	5
M_8	Translog, u_{it} half-normal, firm-specific β	hierarchical	–923.241	4.85×10^{-90}	6
M_9	Translog, u_{it} exponential, firm-specific intercept	hierarchical – TRE model	–717.586	≈ 1	1
M_{10}	Translog, u_{it} half-normal, firm-specific intercept	hierarchical – TRE model	–791.139	1.14×10^{-32}	3

CD – Cobb-Douglas production function; M – model; TRE model – the true random effects model, in which only the intercept varies and other coefficients are fixed between farms
Hierarchical prior means that the coefficients vary between farms (i.e. parameters are firm-specific) and thus in the non-hierarchical model the slope parameters are not individual-specific (i.e. are fixed)
Source: Authors' own calculations based on FADN (2020)

parameters. Both models are based on the exponential distribution for the inefficiency term. Therefore, in the empirical part of the study, we will first present all the results based on model M_9, which will be compared with the results derived from the other specifications.

Since this study's main aim was to assess the level of technical efficiency, its average scores from the competing models are therefore presented in Table 3. The technical efficiency scores obtained can be compared with respect to the following assumptions of the models they are derived from: type of prior for technology parameters (hierarchical *vs.* non-hierarchical), type of production function and type of distribution for the inefficiency term. The technical efficiency in every model with a hierarchical prior (i.e. M_5, ..., M_{10}) is higher than in the corresponding model with a non-hierarchical prior (M_1, ..., M_4). However, the translog models' technical efficiency score is similar, no matter which type of prior, hierarchical or non-hierarchical, was used. Similarly, Čechura (2010) and Emvalomatis (2012) obtained higher technical efficiency scores in random coefficients models. The differences in the mean technical efficiency score due to the form of production function are negligible. Correspondingly, there are virtually no differences in average technical efficiency scores between M_1 and M_3 or M_2 and M_4. Simi-

Table 3. Average firm-level technical efficiency estimates and their changes over time

Year	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
2004	0.86	0.80	0.86	0.80	0.88	0.83	0.87	0.81	0.87	0.82
2005	0.85	0.78	0.84	0.78	0.86	0.80	0.85	0.78	0.85	0.79
2006	0.84	0.78	0.84	0.78	0.86	0.80	0.85	0.78	0.85	0.80
2007	0.86	0.80	0.86	0.80	0.88	0.83	0.87	0.81	0.88	0.83
2008	0.84	0.78	0.84	0.77	0.85	0.80	0.84	0.78	0.85	0.79
2009	0.84	0.78	0.84	0.78	0.86	0.81	0.85	0.79	0.85	0.80
2010	0.85	0.80	0.85	0.79	0.88	0.83	0.86	0.81	0.87	0.82
2011	0.85	0.79	0.84	0.78	0.86	0.81	0.85	0.79	0.86	0.81
Mean	0.85	0.80	0.85	0.79	0.87	0.82	0.86	0.80	0.87	0.81

M – model
Source: Authors' calculations based on FADN (2020)

https://doi.org/10.17221/347/2020-AGRICECON

larly, there is a close resemblance in efficiency scores between M_5 and M_7 or M_6 and M_8.

Considering the distribution of the inefficiency term, it should be noted that the technical efficiency is higher in the models with exponential distribution (in the even-numbered models) than in the ones with half-normal distribution, even though the prior mean of u_{it} is similar between the models. This indicates that estimates of technical efficiency are sensitive to the assumption made regarding the type of distribution of inefficiency. This finding is supported by Marzec and Pisulewski (2019), who also analysed crop farms in Poland, but using a truncated-normal SF model and obtained a substantially lower mean technical efficiency score (63%).

This being said, the computed linear correlation of the efficiency estimates based on M_5 and M_6 or M_9 and M_{10} is about 0.97. Therefore, the ranking of farms is not affected by choice of distribution. Similarly, in the study conducted by Marzec and Pisulewski (2019), the correlation between half-normal and truncated-normal technical efficiency scores was high (0.99). The form of the production function makes as well little difference in the farms' rankings because the geometric mean of the Spearman correlations is equal to 0.95. The distinction between the hierarchical and non-hierarchical structure of the models has a slightly stronger impact on the rankings' differences (the aforementioned mean equals 0.91).

Our findings show that the average posterior mean of the technical efficiency scores was 87% in models M_5 and M_9, and very similar values were obtained for the other models with exponential distribution. This result is similar to that obtained by Čechura et al. (2015), who showed the technical efficiency of Polish cereal farms to be 84%. It is noteworthy that the efficiency scores obtained for M_4 and M_{10} are consistent with the results for Polish crop farms obtained with different methods, i.e. Maximum Simulated Likelihood, presented by Pisulewski and Marzec (2019). Moreover, the latter study showed that models that account for two kinds of inefficiency, i.e. transient and persistent inefficiency (models M_4 and M_6 in their study) lead to a lower overall technical efficiency score.

Figure 1 presents details about the TE scores of Polish crop farms (model M_9). The majority of farms (72%) have an efficiency ranging from 0.85 to 0.95. Furthermore, only 7% of farms' efficiencies are low, i.e. below 0.7. We also find that the median technical efficiency score is around 0.89, and so it is more than the prior assumed, i.e. 0.8. The results obtained from the other models are quite similar, especially including M_5.

In all the considered models, inefficiency is modelled as time-varying. Moreover, it changes over time in a random and unstructured way. Therefore, these models provide no information about the causes or direction of the change in efficiency. Table 3 suggests that there is little variation in technical efficiency scores over time and that there is no trend in changes in TE scores. This result contradicts the findings of Čechura et al. (2015) who showed that, over the period from 2004 to 2011, there were substantial changes in the technical efficiency of Polish cereal farms. Therefore, it is worthwhile testing whether the observed jumps in the efficiency scores from model M_9 are statistically significant. Furthermore, several variables, which can potentially explain the differences in technical efficiency between farms, are defined. These are agricultural policy variables and farms' characteristics. The former variables include investment subsidies and less favoured areas (LFA) subsidies, while the latter consists of specialisation, economic size, land size, non-rented land (use of own land only) and own labour (use of family labour only on their farms). The aforementioned variables are expressed on a binary scale. For example, the two first variables are used to express whether the farm received funding from the EU or not. Moreover, specialisation equals one if crop production in the period t is the main source of income and zero otherwise; while if the value of the total Standard Output (SO) of the farm is not less than EUR 50 000, it is a large unit (economic size is equal to one); otherwise, it is a small one.

Subsequently, we then performed the panel regression of logit transformed efficiency scores, i.e. ln[TE/(1 – TE)], on the aforementioned determinants. Due to the transformation of the dependent var-

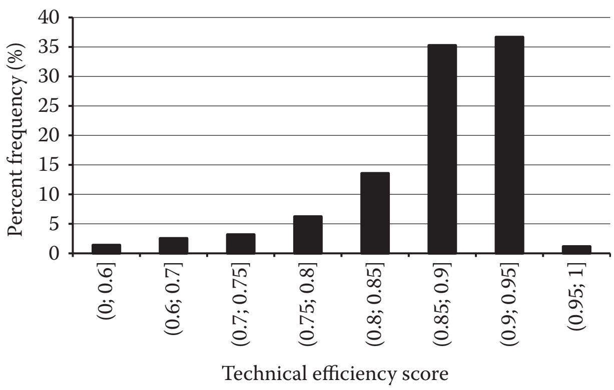

Figure 1. Percentage distribution of estimates of technical efficiency scores (model M_9)

Source: Authors' calculations based on FADN (2020)

https://doi.org/10.17221/347/2020-AGRICECON

Table 4. Standard panel regression of technical efficiency scores on the binary covariates (fixed effects model)

Independent dummy variable with values yes (1) or no (0)	Average value in the sample	Parameter estimate	*P*-value
Specialisation	0.91	–0.16	$\approx 10^{-6}$
Less favoured area (LFA) subsidies	0.29	–0.05	0.1727
Investment subsidies	0.14	–0.13	0.0002
Economic size (large)	0.20	0.13	0.0052
Land area size (large)	0.77	0.21	0.0001
Non-rented land	0.29	–0.07	0.0762
Family labour only	0.50	0.04	0.2110
***F*-test statistic values**			
F-test for individual effects	–	$F(659; 4\,606) = 0.54$	≈ 1
F-test for time effects	–	$F(7; 4\,606) = 19.9$	$\approx 10^{-23}$

Source: Authors' calculations based on FADN (2020)

iable, the dummy variables parameters are interpreted as quasi-elasticities and the positive sign of the coefficients means the associated variables have a negative effect on technical efficiency. Moreover, we considered two models, including individual- and time-specific effects treated as random or fixed. The Hausman test indicates that fixed effects is the appropriate estimator for this regression. The detailed results of this regression model are presented in Table 4.

Furthermore, the results of the *F*-test indicate the statistical significance of the difference in the time-specific effects. Additionally, it was revealed that the *F*-test did not reject the null hypothesis of all individual-specific effects equal to zero. This can be explained by the fact that the TRE model, given by the Equation (2), already includes a heterogeneous intercept α_i, which captures the variation between individuals.

The results summarised in Table 4 indicate that investment subsidies and specialisation in agricultural production have a negative and significant impact on efficiency score. Moreover, farms that use only their own land have higher technical efficiency scores. However, the use of only family labour by farms and receiving LFA subsidies is found to be statistically non-significant. Thus, there are no significant differences in efficiency between these two types of farms. Finally, these findings suggest that larger farms are more efficient than smaller ones. Additionally, among all variables examined in the study, economic size is one of the most important predictors of efficiency.

These results obtained by regressing the estimates of efficiency scores on some additional variables point to some important implications for assumptions about the model [Equation (1)]. Namely, there is a need to take into account the influence of the factors mentioned above in a one-step procedure based on the correctly specified stochastic frontier model. Without a doubt, it will be of interest for future research.

Table 5 contains a summary of information about the production function itself, i.e. in case of a typical crop

Table 5. Production elasticity estimates and returns to scale (RTS) at sample mean for inputs (standard deviation in parentheses)*

Model elasticity	M_2		M_5		M_7		M_9	
Capital	0.001	(± 0.008)	0.028	(± 0.014)	0.017	(± 0.010)	0.029	(± 0.013)
Labour	0.263	(± 0.010)	0.265	(± 0.018)	0.258	(± 0.011)	0.264	(± 0.015)
Materials	0.873	(± 0.013)	0.659	(± 0.021)	0.781	(± 0.015)	0.627	(± 0.018)
Area	0.010	(± 0.009)	0.198	(± 0.018)	0.101	(± 0.011)	0.236	(± 0.015)
RTS	1.148	(± 0.009)	1.150	(± 0.018)	1.157	(± 0.013)	1.156	(± 0.015)
t (trend)	–0.001	(± 0.002)	–0.002	(± 0.002)	0.001	(± 0.002)	–0.0017	(± 0.0015)

M – model; RTS – returns to scale

*Inputs equal to the arithmetic mean of the data on a logarithmic scale

Source: Authors' calculations based on FADN (2020)

https://doi.org/10.17221/347/2020-AGRICECON

farm, i.e. at the geometric mean of the data, all models show that the highest output elasticity is with respect to materials, while the lowest is with respect to capital. The main difference between the models concerns the elasticity with respect to materials and area. In particular, the impact of the former input is higher in the non-hierarchical models, while it is lower in the hierarchical models. In contrast, the influence of the latter input on production is lower in the non-hierarchical models, while it is higher in the hierarchical ones. Furthermore, the results indicate that the basic assumptions underlying production theory are satisfied. It is found that, in model M_9, the sufficient condition for the law of diminishing marginal productivity at the geometric mean of data is fulfilled for all inputs except labour. In the sample of all farms, the majority of farms satisfied this law for all inputs, i.e. 83% for capital, 37% for labour, 100% for materials and 78% for the last input. Similar results were obtained in model M_5 with the CD function with individual-specific parameters, but this condition is weaker for the CD than for the translog function. Additionally, as shown in Table 5, trend parameter is statistically non-significant in all investigated models. Therefore, it implies that there was no technical progress over the whole analysed period (2004–2011).

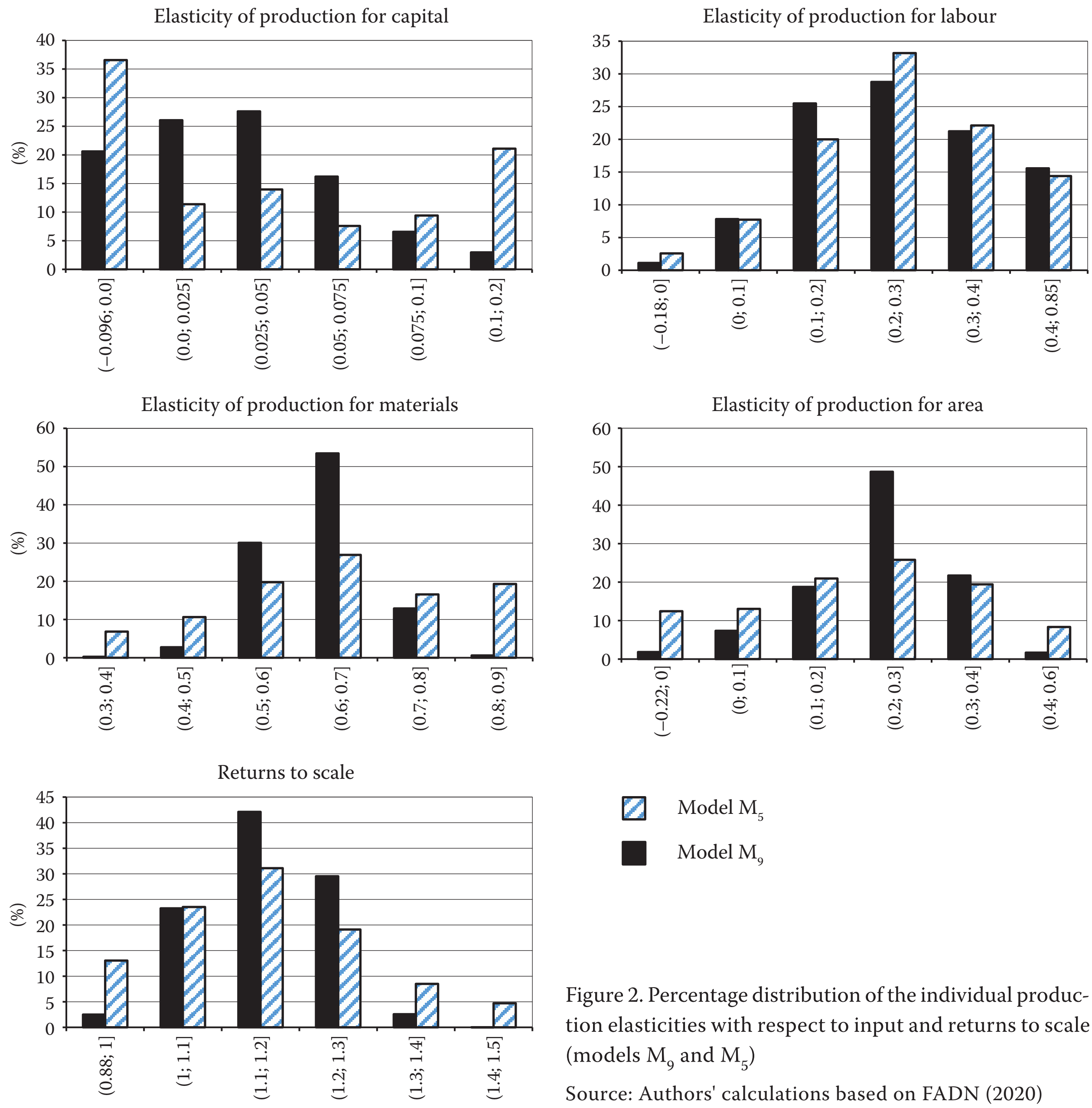

Figure 2. Percentage distribution of the individual production elasticities with respect to input and returns to scale (models M_9 and M_5)

Source: Authors' calculations based on FADN (2020)

https://doi.org/10.17221/347/2020-AGRICECON

In the context of hierarchical modelling, the translog function can be interpreted as the Cobb-Douglas type function with elasticities that depend on the level of inputs in a deterministic way. Based on Figure 2 it can be seen how heterogeneous are farms with respect to each production factor. For example, according to the baseline model M_9, in case of production elasticity with respect to capital, the 21% of farms are characterised by negative elasticity. However, in the case of the majority of farms, i.e. 76% of farms, the elasticity is positive and less than 1. The output elasticity with respect to labour is negative in only 1% of cases, while there are neither negative values nor values above 1 observed in the case of elasticity with respect to materials. In case of elasticity concerning area, 2% of farms are shown to exhibit negative values. The last production characteristic, which can be derived, is the returns to scale. Figure 2 below implies that the majority of Polish crop farms, i.e. 97% of them, operate under increasing returns to scale. Comparing the results from Figure 2 and Table 4, we can observe that the distributions for output elasticities obtained from both models have little difference between them, but their estimates resulting from the CD function (M_5) are more dispersed compared to the translog (M_9) function. In summary, there are noticeable differences between both models on this question, but they are not of sufficiently crucial importance as to require further discussion.

CONCLUSION

In the present study, we have investigated several competing stochastic frontier models using a Bayesian inference. It was revealed that the technical efficiency scores mainly differ between the models with exponential and half-normal distribution for the inefficiency term. The type of prior distribution for the parameters, i.e. non-hierarchical and hierarchical, which correspond to the "fixed" and random coefficients models, does not significantly affect the technical efficiency score, although in the models with a hierarchical prior this score is slightly higher. Moreover, the production elasticities with respect to materials and area slightly vary between the hierarchical and non-hierarchical models. Additionally, no significant differences were noticed due to the form of the production function.

The comparison of models, based on Bayes factors, indicates that, among the models considered, the TRE model is the best. Consequently, the technical efficiency scores obtained from this model were investigated further. In particular, the regression of logit transformed efficiency scores on its determinants, time dummies and farm dummies was performed. It suggests that the differences over time are significant. Moreover, among the analysed determinants, only LFA subsidies and family labour are proved to be statistically non-significant. The impact of specialisation and non-rented land are revealed to be negative, while the other statistically significant determinants have a positive effect on TE. The investigated production characteristics show that the regularity conditions of the production function are satisfied for most of the inputs and in the case of the majority of farms.

REFERENCES

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

Alvarez A., Arias C., Greene W. (2004): Accounting for unobservables in production models: Management and inefficiency. Economic Working Papers No., 72. Seville, Centro de Estudios Andaluces.

Areal F.J., Balcombe K., Tiffin R. (2012): Integrating spatial dependence into stochastic frontier analysis. Australian Journal of Agricultural Resource Economics, 56: 521–541.

Baráth L., Ferto I., Bojnec S. (2018): Are farms in less favoured areas less efficient? Agricultural Economics, 49: 3–12.

Bojnec Š., Latruffe L. (2009): Determinants of technical efficiency of Slovenian farms. Post-Communist Economies, 21: 117–124.

Broeck van den J., Koop G., Osiewalski J., Steel M.F.J. (1994): Stochastic frontier models: A Bayesian perspective. Journal of Econometrics, 61: 273–303.

Caudill S.B. (2003): Estimating a mixture of stochastic frontier regression models via the EM algorithm: A multiproduct cost function application. Empirical Economics, 28: 581–598.

Čechura L. (2010): Estimation of technical efficiency in Czech agriculture with respect to firm heterogeneity. Agricultural Economics – Czech, 56: 183–191.

Čechura L., Hockmann H., Malý M., Kroupová Z. (2015): Comparison of technology and technical efficiency in cereal production among EU countries. Agris on-line Papers in Economics and Informatics, 7: 27–37.

Čechura L., Grau A., Hockman H., Levkovych I., Kroupova Z. (2017): Catching up or falling behind in European agriculture: The case of milk production. Journal of Agricultural Economics, 68: 206–227.

Chib S. (1995): Marginal likelihood from Gibbs output. Journal of the American Statistical Association, 90: 1313–1321.

https://doi.org/10.17221/347/2020-AGRICECON

Emvalomatis G. (2012): Productivity growth in German dairy farming using a flexible modelling approach. Journal of Agricultural Economics, 63: 83–101.

FADN (2020): Polish Farm Accountancy Data Network. [Dataset]. Unpublished Data.

Feng G., Zhang X. (2012): Productivity and efficiency at large community banks in the US: A Bayesian true random effects stochastic distance frontier analysis. Journal of Banking and Finance, 36: 1883–1895.

Greene W.H. (2005): Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics, 126: 269–303.

Greene W.H. (2012): Econometric Analysis. 7th Ed. Essex, England, Pearson Education: 1–1228.

Hildreth C., Houck J.P. (1968): Some estimators for a linear model with random coefficients. Journal of the American Statistical Association, 63: 584–595.

Kalirajan K.P., Obwona M.B. (1994): Frontier production function: The stochastic coefficients approach. Oxford Bulletin of Economics and Statistics, 56: 87–96.

Koop G., Osiewalski J., Steel M. (1997): Bayesian efficiency analysis through individual effects: Hospital cost frontiers. Journal of Econometrics, 76: 77–105.

Latruffe L., Balcombe K., Davidova S., Zawalińska K. (2004): Determinants of technical efficiency of crop and livestock farms in Poland. Applied Economics, 36: 1255–1263.

Marzec J., Pisulewski A. (2017): The effect of CAP subsidies on the technical efficiency of Polish dairy farms. Central European Journal of Economic Modelling and Econometrics, 9: 243–273.

Marzec J., Pisulewski A. (2019): The measurement of time-varying technical efficiency and productivity change in Polish crop farms. German Journal of Agricultural Economics, 68: 15–27.

Njuki E., Bravo-Ureta B.E., O'Donnell C.J. (2019): Decomposing agricultural productivity growth using a random-parameters stochastic production frontier. Empirical Economics, 57: 839–860.

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

Pisulewski A., Marzec J. (2019): Heterogeneity, transient and persistent technical efficiency of Polish crop farms. Spanish Journal of Agricultural Research, 17: 1–14.

Rossi P., Allenby G., McCulloch R. (2005): Bayesian Statistics and Marketing. Chichester, England, John Wiley & Sons: 1–348.

Skevas I., Emvalomatis G., Brümmer B. (2018): Productivity growth measurement and decomposition under a dynamic inefficiency specification: The case of German dairy farms. European Journal of Operational Research, 271: 250–261.

Skevas I. (2019): A hierarchical stochastic frontier model for efficiency measurement under technology heterogeneity. Journal of Quantitative Economics, 17: 513–524.

Swamy P.A. (1970): Efficient inference in a random coefficient regression model. Econometrica, 38: 311–323.

Tsionas E.G. (2002): Stochastic frontier models with random coefficients. Journal of Applied Econometrics, 17: 127–147.

Zhu X., Lansink A.O. (2010): Impact of CAP subsidies on technical efficiency of crop farms in Germany, the Netherlands and Sweden. Journal of Agricultural Economics, 61: 545–564.

Received: August 28, 2020
Accepted: February 1, 2021