ISSN 1822-6760. Management theory and studies for rural business and infrastructure development. 2012. Vol. 34. Nr. 5. Scientific journal

APPLICATION OF THE BOOTSTRAPPED DEA FOR THEANALYSIS OF LITHUANIAN FAMILY FARM EFFICIENCY

Tomas Baležentis, Irena Kriščiukaitienė
Lithuanian Institute of Agrarian Economics

Farming efficiency is one of the most important factors determining profitability and viability in general of the agribusiness. The aim of this paper is to analyze the dynamics of efficiency in Lithuanian family farms by the means of the bootstrapped Data Envelopment Analysis in order to propose certain guidelines for inefficiency mitigation. Stochastic kernels are then employed to estimate densities of the efficiency scores for different farming types. The research covers years 2004–2009 and is based on farm-level Farm Accountancy Data Network (FADN) data. The stochastic kernel for livestock farms exhibited a small range of efficiency scores. Therefore, these farms achieved a higher convergence as well as a higher average TE from the standpoint of the analyzed farming types. The mixed farms, though, were peculiar both a sort of bi-modal distribution of their efficiency scores.

Key words: efficiency, family farms, bootstrapping, data envelopment analysis.
JEL codes: C440, C610, Q100, Q130.

Introduction

The Lithuanian agricultural sector gained a momentum for transformations during Lithuania's integration into the European Union (EU). Therefore, it is important to analyze the recent trends of the sectoral performance and thus draw reasonable policy guidelines. One of the most important indicators describing the performance of an economic sector is its productive efficiency. This paper focuses on farm-level indicators to estimate the efficiency of Lithuanian family farms. Indeed, the efficiency of the family farms is important not only in the economic sense, but it does also influence the viability of the rural areas.

Data Envelopment Analysis (DEA) is one the most celebrated methods employed for productive efficiency analysis (Rimkuvienė, 2010; Bojnec, 2011). However, the latter method is a non-parametric one and thus attributes the whole distance from the efficiency frontier to inefficiency. Accordingly, random error remains ignored. The bootstrapped DEA, fuzzy DEA, and stochastic DEA enable to mitigate this draw-back. This paper focuses on application of the bootstrapped DEA, which was introduced by L. Simar and P. W. Wilson (1998, 2000a, 2000b). A. Assaf and K. M. Matawie (2010) employed the bootstrapped DEA to assess the efficiency of health care foodservice facilities. Aldea and Ciobanu (2011) and Aldea et al. (2012) used the bootstrapped DEA to analyze the renewable energy production efficiency across the EU Member States. G. Halkos and E. Tzeremes (2012) utilized the bootstrapped DEA for the analysis of the Greek renewable energy sector. Noteworthy, the Lithuanian agricultural sector has not been analyzed by the means of the bootstrapped DEA yet. The analysis of efficiency can be carried out by visualising the densities of the efficiency scores (Simar, 2006; Mugera, 2011). This paper, therefore, employs stochastic kernel technique to plot the kernel densities and thus draw the conclusions about the dynamics of efficiency in Lithuanian family farms.

The aim of this paper is to analyze the dynamics of efficiency in Lithuanian family farms by the means of the bootstrapped DEA in order to identify the underlying causes of inefficiency. The research relies on micro data. Specifically, 200 family farms reporting to the Farm Accountancy Data Network (FADN) were chosen for the analysis. The research covers the period of 2004–2009. The following methods were employed for the research: DEA, bootstrapping, and kernel density estimation.

The paper is organized as follows: Section 1 discusses the preliminaries for the bootstrapped DEA. Section 2 focuses on the stochastic kernels. Finally, Section 3 describes the data used and presents the results of analysis.

1. The Bootstrapped DEA

Given DEA is a non-parametric technique, it does not allow for a statistical noise in estimations of the efficiency measures. Specifically, the efficiency scores are calculated by employing linear programming models rather than estimated. The ordinary DEA, hence, is unable to yield the information on sensitivity of the obtained results. It was L. Simar and P. W. Wilson (1998) who presented a solution for the latter problem, namely the bootstrapping approach for DEA. The underlying idea is to estimate the population distribution of the DEA efficiency scores, thus making it possible to perform hypothesis testing on the efficiency scores (Assaf, 2010).

Basically, the definition of bootstrapping encompasses iterative random sampling from the observed sample data. If the procedure is repeated thousands of times, one can derive the pseudo estimates from these samples. The latter pseudo estimates, in turn, define an empirical distribution related to the estimator of interest. This distribution is an approximation of the true underlying sampling distribution of the estimator. A. Assaf and K. Matawie (2010) presents the following outline of bootstrapping procedure: Considering, for example, a random sample $X = (X_1, X_2, \dots, X_n)$ from a population with unknown distribution function *F*, the objective is to estimate the sampling distribution of an distribution function of some pre-defined random variable $R(X, F)$, using a real data set, *x*, where $x = (x_1, x_2, \dots, x_n)$ represents the observed realization of $X = (X_1, X_2, \dots, X_n)$. To be specific, the bootstrapping procedure begins with construction of a sample probability distribution $\hat{F}$ by assigning probability of $1/n$ at each point in the observed sample, $x_1, x_2, \dots, x_n$. A random sample is then drawn with replacement from $\hat{F}$ while $\hat{F}$ is fixed at its observed value. The obtained sample $X^* = (X_1^*, X_2^*, \dots, X_n^*)$ is defined as the bootstrap sample $X_i^* = x_i^*$, $x_i^* \overset{ind}{\sim} \hat{F}$, $i = 1, 2, \dots, n$. To cap it all, the distribution of the random sample $R(X, F)$ is approximated by the bootstrap distribution of $R^* = R(X^*, \hat{F})$.

As it was already said, L. Simar and P. Wilson (1998, 2000a, 2000b) presented a bootstrapping algorithm for input-oriented DEA model. This model, however, can be generalized to facilitate output-oriented measures. A production set, defined as

$$T = \left\{(x, y) \in R_{+}^{m+n} \middle| x \text{ can produce } y\right\}, \tag{1}$$

relates the amount of *m* inputs, *x*, to respective amount of *n* outputs, *y*. For a given output level, *y*, the set of inputs that is feasible in terms of the technology is defined as the correspondence:

$$X(y) = \left\{x \in R_{+}^{m} \mid (x, y) \in T\right\}. \tag{2}$$

The efficient production frontier is then defined as a subset of $X(y)$:

$$X_e(y) = \left\{x \mid x \in X(y), \theta x \notin X(y), \forall \in (0,1)\right\}, \tag{3}$$

so that input–output bundles of $X_e(y)$ cannot remain feasible in case of a further contraction of inputs at a given level of outputs. The efficiency score, θ is an input-oriented Farrel efficiency measure: $\theta_k = \min\{\theta \mid \theta x_k \in X(y_k)\}$, where *k* denotes a respective DMU.

In the real world the sets *T*, $X(y)$, and $X_e(y)$ are unknown. Consequently the efficiency score, θ_k, is also unknown. Indeed, it is assumed that a certain data-generating process generates a random sample, $X = \left\{(x_k, y_k) \mid k = 1, 2, ..., K\right\}$, of *K* homogeneous firms (DMUs). This sample defines, by the virtue of some method the equivalents of $\widehat{X}(y)$, and $\hat{X}_e(y)$, and $\hat{\theta}$. Indeed, $\widehat{\theta}_k$ can be obtained by employing a DEA model:

$$\widehat{\theta}_k = \min\left\{\theta \mid \sum_{k=1}^{K} \lambda_k x_k \le \theta x_k, \sum_{k=1}^{K} \lambda_k y_k \ge y_k, \sum_{k=1}^{K} \lambda_k = 1\right\}. \tag{4}$$

The approximate efficiency scores can then be used in the bootstrapping procedure to obtain pseudo-samples of the efficient input vectors:

$$\hat{x}^e(x_k \mid y_k) = \hat{\theta}_k x_k, \tag{5}$$

where $\hat{x}^e(x_k \mid y_k)$ denotes the input level a firm should achieve in order to be on the production frontier. Specifically, a random sample with replacement, θ_k^*, $k = 1, 2, \ldots, K$, is chosen from $(\hat{\theta}_1, \theta_2, \ldots, \theta_K)$ and then the bootstrap inputs are given by

$$x_k^* = \frac{\hat{\theta}_k}{\theta_k^*} x_k. \tag{6}$$

The DEA model is employed to obtain the bootstrap estimates of the efficiency scores, $\hat{\theta}_k^*$, based on the bootstrapped inputs (cf. Eq. 6). The same procedure is iterated *B* times in order to obtain the sampling distribution for θ_k, which, in turn, is used to estimate the bias and to conduct inference on the efficiency scores. It is, however, the very nature of the DEA efficiency scores, that makes the bootstrapping process somehow complicated; for the empirical distribution $\hat{F}$ of $\hat{\theta}$ provides inconsistent estimates of the true density function, *F*. Indeed, the efficiency scores lie in the interval between 0 and 1 and thus make the empirical distribution discontinuous at this interval. The latter issue causes inconsistency of the bootstrap measure. It is due to L. Simar and P. Wilson (1998) that the smoothed bootstrapping procedure can be employed to obtain the consistent estimates. Specifically, a

Gaussian kernel density estimator is used to obtain $\hat{F}$ alongside with the reflection method (Silverman, 1986) to tackle the problem that F is truncated at 1.

It is due to L. Simar and P. Wilson (1998) that the Shepard efficiency measures are more suitable for the computations involved in the bootstrapping procedure. Therefore, let $\delta(x,y) \equiv \theta^{-1}(x,y)$ be the Shepard efficiency measure. L. Simar and P. Wilson (2008) presented the following algorithm of bootstrapping for DEA:

1. From the original data set, X_K, compute $\hat{\delta}_k = \hat{\delta}(x_k, y_k), \forall k = 1,2,...,K$.
2. Select a value for bandwidth h.
3. Generate $\beta_1^*, \beta_2^*, ..., \beta_K^*$ by drawing with replacement from the set $\{\hat{\delta}_1, \hat{\delta}_2, ..., \hat{\delta}_K, (2-\hat{\delta}_1), (2-\hat{\delta}_2), ..., (2-\hat{\delta}_K)\}$.
4. Draw ε_k^* independently from the kernel function, $K(\cdot)$, and compute $\beta_k^{**} = \beta_k^* + h\varepsilon_k^*$ for each $k = 1,2,\ldots,K$.
5. For each $k = 1,2,\ldots,K$, compute $\beta_k^{***} = \bar{\beta}^* + \frac{\beta_k^{**} - \bar{\beta}^*}{(1 + h^2\sigma_K^2\sigma_\beta^{-2})^{1/2}}$, where $\bar{\beta}^* = K^{-1}\sum_{k=1}^{K}\beta_k^*$ is the sample mean of the β_k^*, $\sigma_\beta^2 = K^{-1}\sum_{k=1}^{K}(\beta_k^* - \bar{\beta}^*)^2$ is the sample variance of the β_k^*, and σ_K^2 is the variance of the probability density function used for the kernel function; and then compute $d_k^* = \begin{cases} 2 - \beta_k^{***}, & \forall \beta_k^{***} < 1 \\ \beta_k^{***}, & \text{otherwise} \end{cases}$.
6. Define the bootstrap sample $X_K^* = \{(x_k^*, y_k) \mid k = 1,2,\ldots,K\}$, where $x_k^* = \delta_k^* \hat{x}^e(y_k) = \delta_k^* \hat{\delta}_k^{-1} x_k$.
7. Compute the DEA efficiency estimates $\hat{\delta}^*(x_k, y_k)$ for the fixed point (x_k, y_k) from the original data set using X_K^* as a reference set.
8. Finally, steps 1–7 are iterated B times ($B \geq 2000$) to obtain a set of bootstrap estimates $\hat{\delta}_b^*(x,y) \mid b = 1,2,\ldots,B$.

The mean of a bootstrap estimator is then used as an approximation of the original VRS DEA estimator, $\hat{\hat{\delta}}_{VRS}(x,y)$:

$$bias_B\left(\hat{\theta}_{VRS}(x,y)\right) = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_{VRS,b}^*(x,y) - \hat{\theta}_{VRS}(x,y). \tag{7}$$

An additional factor is required for the both right hand side terms in Eq. 7 in order to correct for the effects of different sample sizes in the true world and bootstrap world (Simar, Wilson, 2008), viz. $(m/n)^{(2/(N+M+1))}$, where n and m are the sizes of the original and bootstrap data sample, respectively; N and M are the numbers of inputs and outputs, respectively.

The bias-corrected estimator of the DEA efficiency score, $\hat{\hat{\theta}}_{VRS}(x,y)$, is then given by

$$\begin{aligned}\hat{\hat{\theta}}_{VRS}(x,y) &= \hat{\theta}_{VRS}(x,y) - bias_B\left(\hat{\theta}_{VRS}(x,y)\right) \\ &= 2\hat{\theta}_{VRS}(x,y) - \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_{VRS,b}^*(x,y)\end{aligned}. \tag{8}$$

Fig. 1 depicts the principle of the bootstrapped DEA. One can define the initial efficiency frontier based on technology set $\hat{T}$ and thus observe the input-oriented technical efficiency score $\hat{\theta}$. The bootstrap procedure then approximates the unobserved real efficiency frontier *T*.

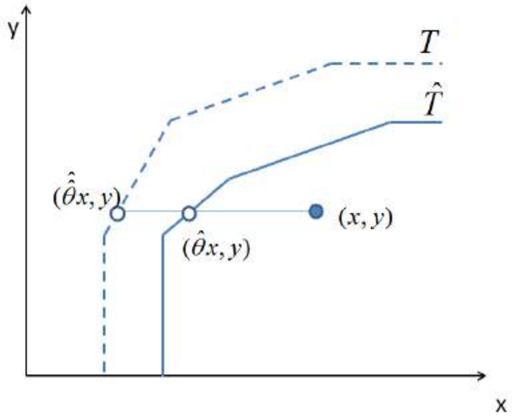

Fig. 1. The bootstrapped input-oriented DEA model

Given the nature of DEA, the efficiency frontier can move outwards as a result of the bootstrapping procedure, but never inwards. This is because the efficiency frontier is drawn on a basis of DMUs peculiar with the highest productivity. Therefore, only those bootstrap samples with input–output bundles related to a higher productivity are to push the approximated (real) production frontier. Accordingly, the bias-corrected efficiency scores, $\hat{\hat{\theta}}$, are lower or equal to those obtained from the original data sample.

2. Kernel smoothing and density estimation

The obtained efficiency scores can be visualised in several ways. The most common technique is that of histograms. However, drawing a histogram involves an arbitrary selection of a number of bins, which may lead to biased interpretations. Non-parametric kernel density estimation techniques allow to overcome these issues and plot the underlying density function of a variable (e. g. efficiency score).

As A. Mugera and M. Langmeier (2011) pointed out kernel-based density functions are becoming an increasingly popular tool for efficiency analysis. The main advantage of the kernel densities is smoother density estimates and independence on the number of bins and their width. Furthermore, no distributional assumptions are imposed on the efficiency scores. It is due to L. Simar and V. Zelenyuk (2006) that the three issues have the major importance on kernel density estimation: (i) the underlying variable must have a bounded support, (ii) only the consistent estimates of the efficiency scores are used, and (iii) the underlying value must not violate the continuity assumption. It is Silverman reflection that satisfies conditions (i) and (iii), whereas bootstrapped DEA provides one with consistent estimates as stipulated by (ii).

The two types of density functions can be obtained, namely conditional and unconditional density functions. Conditional functions yield conditional densities that describe the occurrence of a certain variable's values given the values of another variable. The constructed kernels enable to reveal non-linear relationships between the variables. The unconditional density functions based on kernel estimates allow to display non uni-modal distributions.

Following J. S. Racine (2008), let $f(\cdot)$ and $\mu(\cdot)$ be the joint and marginal densities of (X,Y) and *X*, respectively. Let *Y* and *X* be the dependent and independent variables, respectively. Then the stochastic kernel (or the conditional distribution function) can be estimated as

$$\hat{g}(y \mid x) = \frac{\hat{f}(x,y)}{\hat{f}(x)}. \tag{9}$$

A variety of kernels can be employed as the distribution functions. The product Gaussian kernel approximates $\hat{f}(x,y)$ as

$$\hat{f}(x,y) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_x\sqrt{2\pi}} e^{-0.5\left(\frac{x-x_i}{h_x}\right)^2} \frac{1}{h_y\sqrt{2\pi}} e^{-0.5\left(\frac{y-y_i}{h_y}\right)^2}, \tag{10}$$

and $\hat{f}(x)$ as

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_x\sqrt{2\pi}} e^{-0.5\left(\frac{x-x_i}{h_x}\right)^2}, \tag{11}$$

where h_x and h_y are respective bandwidths.

The analysed variables can be both cardinal and ordinal ones. The ordinal variables are treated by employing special kernel functions (Racine, 2008).

3. Data and Results

The technical efficiency (TE) was assessed in terms of the input and output indicators commonly employed for agricultural efficiency and productivity analyses. More specifically, the utilized agricultural area (UAA) in hectares was chosen as land input variable, annual work units (AWU) – as labour input variable, intermediate consumption in Litas, and total assets in Litas as a capital factor. On the other hand, the three output indicators represent crop, livestock, and other outputs in Litas, respectively. Indeed, the three output indicators enable to tackle the heterogeneity of production technology across different farms.

The data for 200 farms selected from the FADN sample cover the period of 2004–2009. Thus a balanced panel of 1200 observations is employed for analysis. The analyzed sample covers relatively large farms (mean UAA – 244 ha). As for labour force, the average was 3.6 AWU. The data were analyzed in a cross–section way.

In order to quantify the change in productivity across different farming types, the farms were classified into the three groups in terms of their specialization. Specifically, farms peculiar with crop output larger than 2/3 of the total output were considered as specialized crop farms, whereas those specific with livestock output larger than 2/3 of the total output were classified as specialized livestock farms. The remaining farms fell into the mixed farming category.

The efficiency scores were obtained by employing the output-oriented bootstrapped DEA model under VRS assumption (*B*=2000). The *FEAR* package (Wilson, 2010) was applied to implement the latter model. The reported efficiency scores are Shepard measures. As Fig. 2 suggests, the difference between the original and the bootstrapped DEA scores was not a decisive one, i. e. 3 p. p. on an average. The highest difference was observed for years 2004, 2007, and 2008, which implies that the highest data variability occurs during technological expansion. This finding might indicate that some farms tend to increase their output during the favourable periods in terms of climatic conditions to a higher extent than the remaining ones. Therefore, there is a need for the further researches into sources and factors of convergence between the Lithuanian family farms from the viewpoint of their productivity and efficiency.

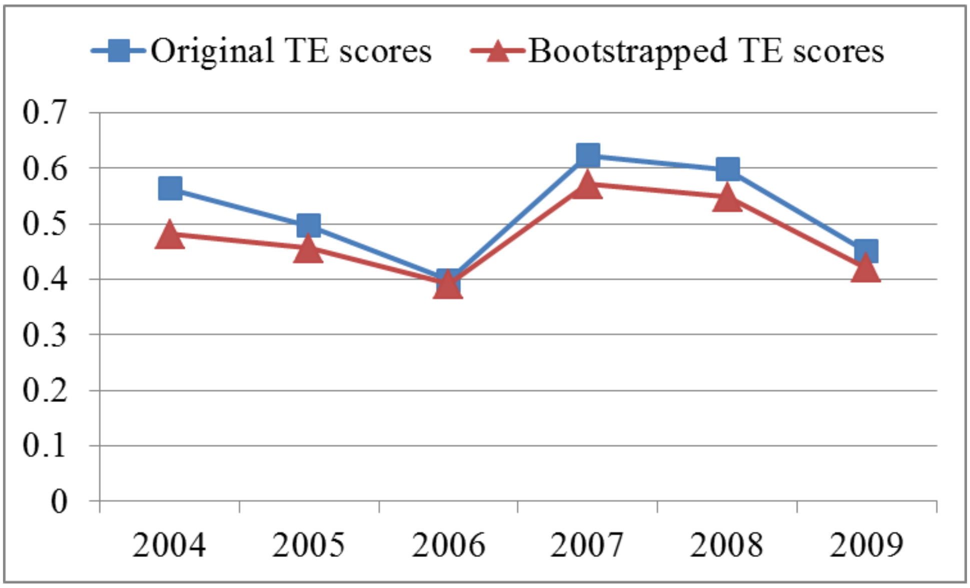

Fig. 2. Technical efficiency (TE) scores obtained by different techniques, 2004–2009

Table 1 further explores the dynamics of the DEA efficiency scores. As one can note, the highest discrepancy between the original TE scores and the bootstrapped ones was observed for the livestock farms (some 14 p. p.). Noteworthy, these discrepancies increased in years 2005 and 2008 to the highest extent. One can therefore assume that the livestock farms are specific with a lagged response to changes in crop markets.

The bootstrapped DEA efficiency scores imply that an average farm should have increased its outputs by a factor of 2.1 (=1/0.48) given the input quantities remain fixed. The same factor was 2.2, 1.7 and 2.1 for crop, livestock, and mixed farms, respectively. The Table 1, however, presents only averages of the estimates. We will further employ stochastic kernels to analyze the distribution of efficiency scores.

Table 1. Average technical efficiency (TE) scores across different farming types, 2004–2009

Year	Original TE scores				Bootstrapped TE scores			
	Crop	Livestock	Mixed	Average	Crop	Livestock	Mixed	Average
2004	0.54	0.67	0.64	0.56	0.49	0.54	0.41	0.48
2005	0.45	0.77	0.59	0.50	0.42	0.58	0.51	0.46
2006	0.35	0.67	0.50	0.40	0.35	0.57	0.43	0.39
2007	0.60	0.83	0.63	0.62	0.56	0.67	0.55	0.57
2008	0.57	0.82	0.60	0.60	0.55	0.62	0.51	0.55
2009	0.43	0.62	0.48	0.45	0.41	0.50	0.41	0.42
Average	0.48	0.72	0.56	0.51	0.46	0.58	0.47	0.48

The stochastic kernels were obtained by employing package *np* (Hayfield, 2008). Specifically, the conditional distribution functions were estimated with years treated as an ordered independent variable and efficiency scores as a continuous numerical variable. Figs. 3–6 present the stochastic kernels for each farming type during 2004–2009.

The density function depicted in Fig. 3 indicates that a certain convergence in efficiency had been achieved during 2004–2009. Specifically, year 2004 was specific with a wide range of efficiency scores with highest densities at TE levels of 0.2–0.3 and 0.4–0.5. Year 2007 was that of increasing efficiency. Therefore, a group of farms emerged with efficiency scores scattered around 0.7. In year 2008, the family farms diverged in their efficiency, whereas year 2009 was that of significant decrease in TE.

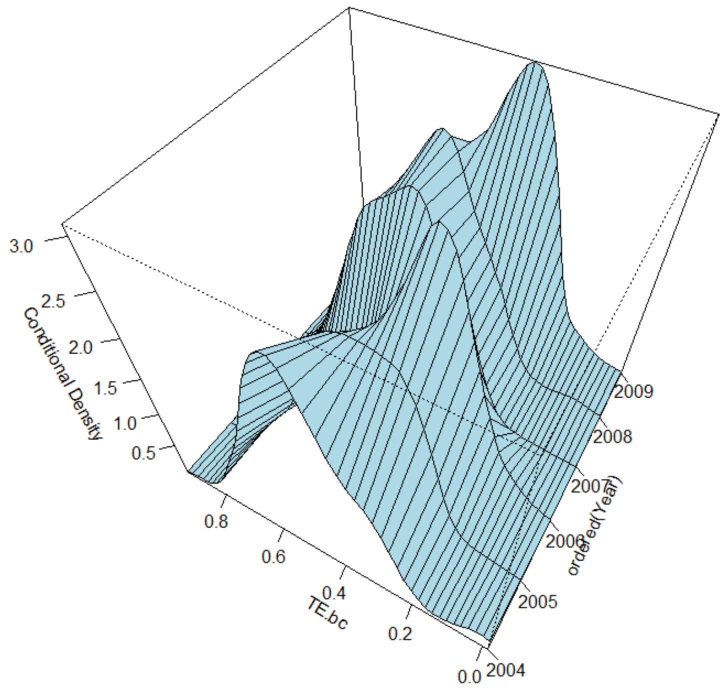

Fig. 3. Stochastic kernel of efficiency scores (all farming types)

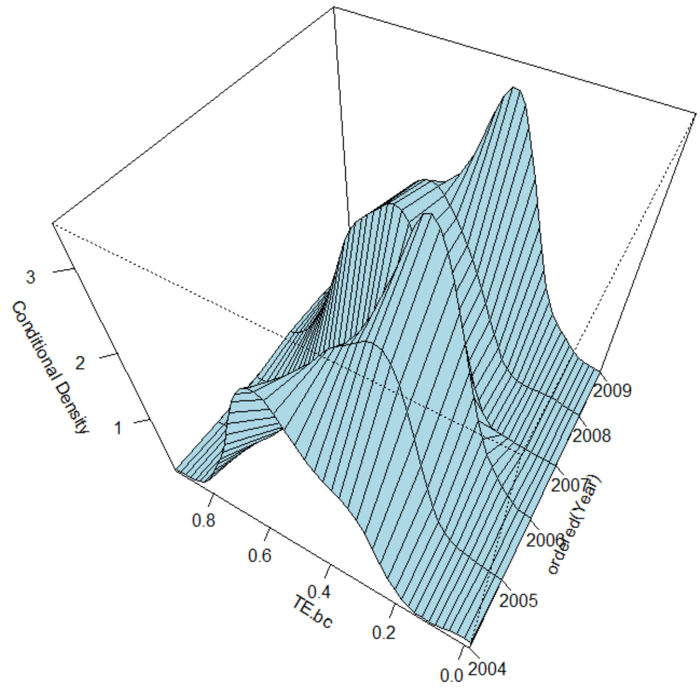

Fig. 4. Stochastic kernel of efficiency scores for crop farms

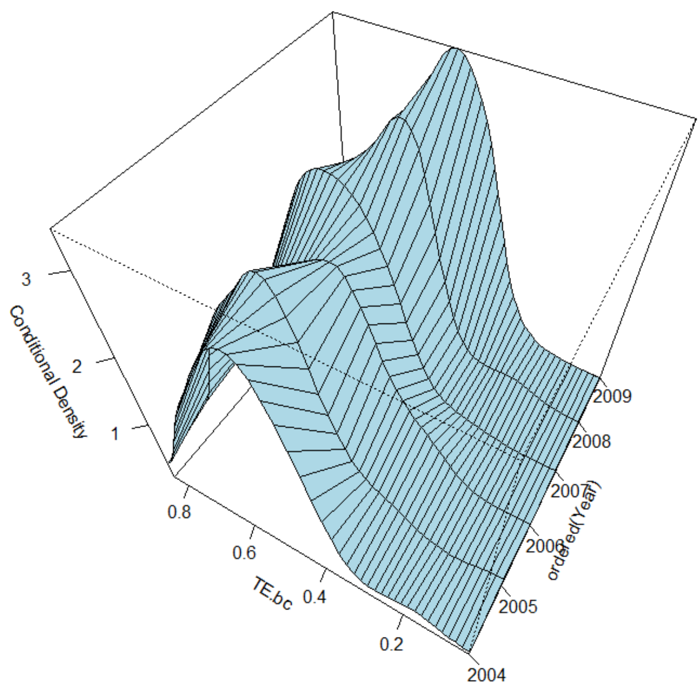

Fig. 5. Stochastic kernel of efficiency scores for livestock farms

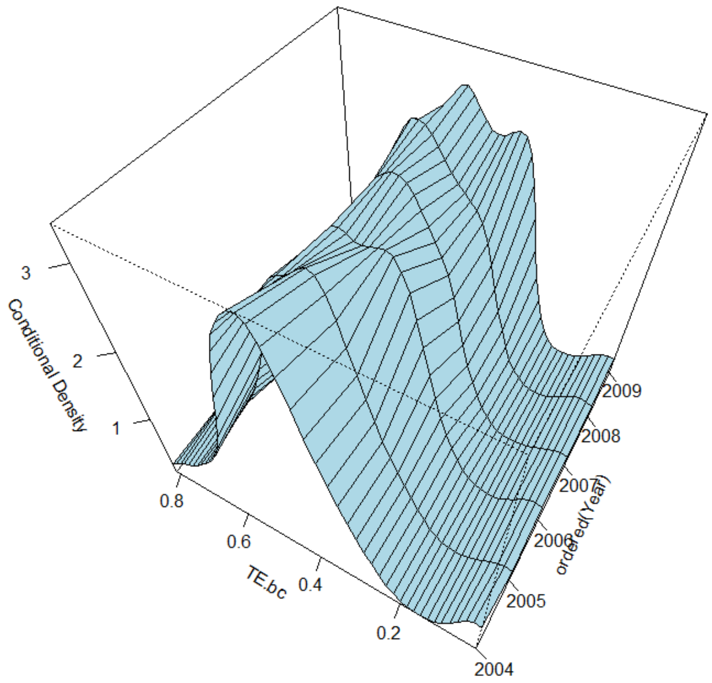

Fig. 6. Stochastic kernel of efficiency scores for mixed farms

Thanks to a high number of crop farms in the sample, kernels exhibited in Figs. 3 and 4 virtually coincide. It is evident that the number of low efficiency crop farms decreased during years 2007–2009, whereas other farming types already had extremely low shares of suchlike under-performing farms.

The stochastic kernel for livestock farms (Fig. 5) exhibits a small range of efficiency scores. Therefore, these farms achieved a higher convergence as well as a higher average TE of some 60 % if compared to the remaining farming types. Furthermore, the share of livestock farms specific with extremely low TE (that below 30 %) was extremely low if compared to the remaining farming types.

The efficiency scores of the mixed farms followed a bi-modal distribution (Fig. 6) with increasing range. In addition, there is a group of mixed farms experiencing extremely low efficiency. Anyway, the average TE score for the mixed farms remained unchanged throughout the whole research period, whereas those of the remaining farming types did decrease.

Conclusions

The bootstrapped DEA efficiency scores imply that an average farm should have increased its outputs twofold given the input quantities remain fixed. The same factor was lower (i. e. 1.7) for the livestock farms.

The stochastic kernel for livestock farms exhibited a small range of efficiency scores. Therefore, these farms achieved a higher convergence as well as a higher average TE from the standpoint of the analyzed farming types. The mixed farms, though, were peculiar with a sort of bi-modal distribution of their efficiency scores. The latter finding implies that future researches should attempt to identify these two farm clusters as well as certain structural measures to increase productive efficiency of the lower performance cluster.

The results of the analysis stress the need for the Lithuanian agricultural policy measures aimed at the convergence of technical efficiency across the different farming types. An especial attention should be paid for crop farm efficiency. Indeed, the crop farming would remain non-profitable in case the current rates of yield were maintained.

The carried out research focused on efficiency score distributions and did not addressed other inter-related issues. Specifically, the further studies could attempt to analyse the relationships between inputs or environmental variables and efficiency scores. Particularly, the non-parametric regression can be employed for these purposes.

References

1. Aldea, A., Ciobanu, A. (2011). Analysis of renewable energy development using bootstrap efficiency estimates // Economic Computation and Economic Cybernetics Studies and Research. Vol. 45. No. 1.
2. Aldea, A., Ciobanu, A., Stancu, I. (2012). The renewable energy development: a nonparametric efficiency analysis // Romanian Journal of Economic Forecasting. Vol. 15. No. 1.
3. Assaf, A., Matawie, K. M. (2010): Improving the accuracy of DEA efficiency analysis: a bootstrap application to the health care foodservice industry // Applied Economics. Vol. 42. No. 27.
4. Bojnec, S., Latruffe, L. (2011). Farm Size and Efficiency during Transition: Insights from Slovenian Farms // Transformations in Business and Economics. Vol. 10. No. 3.
5. Halkos, G. E., Tzeremes. (2012). Analyzing the Greek renewable energy sector: A Data Envelopment Analysis approach // Renewable and Sustainable Energy Reviews. Vol. 16.
6. Hayfield, T., Racine, J. S. (2008). Nonparametric Econometrics: The np Package // Journal of Statistical Software. Vol. 27. No. 5.
7. Mugera, A., Langemeier, M. R. (2011). Does Farm Size and Specialization Matter for Productive Efficiency? Results from Kansas // Journal of Agricultural and Applied Economics. Vol. 43. No. 4.
8. Racine, J. S. (2008). Nonparametric Econometrics: A Primer // Foundations and Trends in Econometrics. Vol. 3. No. 1.
9. Rimkuvienė, D., Laurinavičienė, N., Laurinavičius, J. (2010). ES šalių žemės ūkio efektyvumo įvertinimas // LŽŪU mokslo darbai. Nr. 87(40).
10. Silverman, B. W. (1986). Density Estimation. – London: Chapman and Hall.
11. Simar, L., Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models // Management Science. Vol. 44.
12. Simar, L., Wilson, P. W. (2000a). Statistical inference in nonparametric frontier models: The state of the art // Journal of Productivity Analysis. Vol. 13.
13. Simar, L., Wilson, P. W. (2000b). A general methodology for bootstrapping in nonparametric frontier models // Journal of Applied Statistics. Vol. 27.
14. Simar, L., Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. // Fried, H. O., Lovell, C. A. K., Schmidt, S. S. (Eds.). The Measurement of Productive Efficiency, 2nd ed. – Oxford: Oxford University Press.
15. Simar, L., Zelenyuk, V. (2006). On Testing Equality of Distributions of Technical Efficiency Scores // Econometric Reviews. Vol. 25. No. 4.

BUTSTREPO DUOMENŲ APGAUBTIES ANALIZĖS TAIKYMAS VERTINANT LIETUVOS ŪKININKŲ ŪKIŲ VEIKLOS EFEKTYVUMĄ

Tomas Baležentis, Irena Kriščiukaitienė
Lietuvos agrarinės ekonomikos institutas

Santrauka

Ūkininkavimo efektyvumas yra vienas iš svarbiausių veiksnių, lemiančių agroverslo pelningumą ir gyvybingumą. Šio straipsnio tikslas – pritaikius butstrepo duomenų apgaubties analizės metodą, ištirti Lietuvos ūkininkų ūkių efektyvumo dinamiką ir pasiūlyti neefektyvumo mažinimo priemones. Stochastinių branduolių metodas buvo pritaikytas įvertinant efektyvumo rodiklių skirstinius, būdingus atitinkamiems ūkininkavimo tipams. Tyrimo periodas apima 2004–2009 m. ir remiasi Ūkių apskaitos duomenų tinklo duomenimis. Stochastinis efektyvumo įverčių branduolys gyvulininkystės ūkiams parodė, kad šių ūkių efektyvumo įverčiai yra išsidėstę kompaktiškai (siaurame intervale) ir jiems yra būdinga aukštesnė vidurkio reikšmė, lyginant su kitais nagrinėtais ūkininkavimo tipais. Mišrių ūkių efektyvumo įverčių skirstinys buvo bimodalinis, taigi galima nagrinėti bent du šių ūkių klasterius.

Raktiniai žodžiai: efektyvumas, ūkininkų ūkiai, butstrepas, duomenų apgaubties analizė.
JEL kodai: C440, C610, Q100, Q130.