ANALYSIS OF THE RELATIVE EFFICIENCY OF LITHUANIAN FAMILY FARMS UNDER UNCERTAINTY: ALPHA-QUANTILE FRONTIERS

Tomas Baležentis¹, Rasa Melnikienė², Mette Asmild³, Jens Leth Hougaard⁴, Aistė Galnaitytė⁵

¹ Junior Research Fellow at the Lithuanian Institute of Agrarian Economics. Address: Kudirkos Str. 18/2, 03105 Vilnius, Lithuania. Tel. +370 52 621078. E-mail tomas@laei.lt
² PhD, E-mail laei@laei.lt, ³ Prof., PhD, E-mail jlh@ifro.ku.dk, ⁴ Prof., PhD, E-mail meas@ifro.ku.dk, ⁵ Researcher. E-mail aiste@laei.lt

© Aleksandras Stulginskis University, © Lithuanian Institute of Agrarian Economics

Received 11 11 2013; accepted 10 12 2013

Estimation of the farming efficiency constitutes an important issue related to the policy analysis. However, the efficiency of Lithuanian family farms has not been analysed by the means of the order–alpha frontiers yet. The latter technique enables to estimate the level of efficiency along with sensitivity analysis. This study attempted to analyse the Lithuanian family farm efficiency as well as the associated uncertainties. The probabilistic analysis of the efficiency of Lithuanian family farms suggested that the livestock farming was the most efficient farming type. Even though livestock farming appeared to be a relatively efficient farming type, the number of livestock is decreasing Lithuania. Such changes might be fuelled by both economic and social developments in Lithuania. Therefore, the appropriate measures aimed at fostering livestock farming in Lithuania would contribute to increase in agricultural efficiency and productivity.

Keywords: efficiency, family farms, order–alpha frontier, activity analysis.
JEL Codes: C14, C44, D24, Q12.

1. Introduction

Estimation of the farming efficiency constitutes an important issue related to the policy analysis. However, the agricultural sector features a certain degree of uncertainty, which stems from the varying climatic conditions as well as other environmental factors. As for the farm–level data these are also subject to the sampling bias. Therefore, the research methodology needs to be extended to tackle the aforementioned issues.

Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA) are the two deterministic methods commonly employed for the productivity and efficiency analysis. These methods are quite appealing in the sense that they do not require the explicit assumptions on the functional form of the production function or the error term. However, they do not allow for the statistical noise. The aforementioned deterministic methods therefore were extended to the probabilistic environment. The chance constrained DEA seek to tackle the statistical noise which affects all the observations (Land, 1993; Huang, 2001). Another remedy to the uncertainty in the efficiency analysis is the partial frontier measures (Daraio, 2007). C. Cazals et al. (2002) introduced
the order–\( m \) frontiers which are based on the \( m \) observations randomly drawn from the observed sample to serve as benchmarks. On the other hand, the order–\( \alpha \) frontiers were introduced to define the benchmark by rather setting the probability of dominance, \( \alpha \). It was Y. Aragon et al. (2005) who introduced the concept of the order–\( \alpha \) frontiers in a (partially) univariate framework. A. Daouia and L. Simar (2007) further developed the latter concept by allowing for the multivariate analysis. D. C. Wheelock and P. W. Wilson (2008) offered an unconditional measure of the \( \alpha \text{–} \)efficiency.

The efficiency of Lithuanian family farms has not been analysed by the means of the order–\( \alpha \) frontiers yet. Therefore, there is a need for estimation of efficiency of the Lithuanian agricultural sector based on methodology of the partial frontiers. Indeed, application of multiple techniques increases the robustness of the research. This study, therefore, attempts to analyse the uncertainties associated with the efficiency of Lithuanian family farms. The farm–level data from Farm Accountancy Data Network (FADN) were thus analysed by employing the order–\( \alpha \) efficiency measures.

2. Probabilistic production technology

The activity analysis (Debreu, 1951) defines the production technology by treating the sets of inputs, \( x \in \mathbb{R}^p \), and outputs, \( y \in \mathbb{R}^q \), across the decision making units (DMUs). The technology set, \( T \), consists of all feasible production plans:

\[
T = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^q \mid x \text{ can produce } y\}. \tag{1}
\]

Furthermore, the free disposability of inputs and outputs is assumed (Shepard, 1970), i.e. \( (x, y) \in T \Rightarrow (x', y') \in T \text{ for } x \leq x', y' \leq y \). Note that inequalities between vectors are to be read element-wise throughout the paper.

The input- and output-oriented Farrell (1957) measures of efficiency can be defined, respectively, as:

\[
\theta(x, y) = \inf \{\theta \mid (\theta x, y) \in T \}, \text{ and} \tag{2}
\]

\[
\lambda(x, y) = \sup \{\lambda \mid (x, \lambda y) \in T \}. \tag{3}
\]

The variables \( \theta \in [0, 1] \) and \( \lambda \in [1, +\infty) \) are the input– and output–oriented efficiency scores, respectively. These scores indicate the degree of the proportional contraction (augmentation) of inputs (outputs). The efficient points feature efficiency scores equal to unity. The latter measures render the efficient observations \( (x^\theta(y), y) \in T \), where \( x^\theta(y) = \theta(x, y) x \), for the input direction and \( (x, y^\lambda(x)) \in T \), where \( y^\lambda(x) = \lambda(x, y) y \), for the output direction.

In empirical studies, the set \( T \) and hence the efficiency scores are unknown (Daraio, 2005). Indeed, the quantities of interest are estimated from a random sample of the DMUs, \( \chi_k = \{(x_k, y_k) \mid k = 1, 2, \ldots, K\} \). The non-parametric methods (Farrel, 1957; Charnes, 1978) have been widely employed for efficiency analysis for they are devoid of the over-restrictive hypotheses on the DGP.
In this spirit, a certain DMU, \((x_k, y_k)\), defines an associated production possibility set, \(\tau(x_k, y_k)\), which, under the free disposability of inputs and outputs, can be given as:

\[
\tau(x_k, y_k) = \{(x, y) \in \mathbb{R}_+^{pq} | x_k \leq x, y_k \geq y\}.
\] (4)

The union of the individual production possibility sets (Eq. 4) results in the Free Disposal Hull (FDH) estimator of the underlying technology set, \(T\):

\[
\hat{T}_{FDH} = \bigcup_{k=1}^{K} \tau(x_k, y_k) = \{(x, y) \in \mathbb{R}_+^{pq} | x_k \leq x, y_k \geq y, k = 1, \ldots, K\}.
\] (5)

The efficiency scores can then be obtained by plugging Eq. 5 into Eqs. 2-3.

C. Cazals et al. (2002) and later on C. Daraio & L. Simar (2005) introduced the probabilistic description of the production process. The latter approach is of particular usefulness for estimation of the robust frontiers. The production process, thus, can be described in terms of the joint probability measure, \((X,Y)\) on \(\mathbb{R}_+^{pq} \times \mathbb{R}_+^{pq}\). This joint probability measure is completely characterized by the knowledge of the probability function \(H_{XY}(\cdot, \cdot)\) defined as:

\[
H_{XY}(x, y) = \Pr(X \leq x, Y \geq y).
\] (6)

The support of \(H_{XY}(\cdot, \cdot)\) is \(T\) and \(H_{XY}(x, y)\) can be interpreted as the probability for a DMU operating at \((x, y)\) to be dominated. Note that this function is a non-standard one, with a cumulative distribution form for \(X\) and a survival form for \(Y\).

In the input orientation, it is useful to decompose the joint probability as follows:

\[
H_{XY}(x, y) = \Pr(X \leq x | Y \geq y) \Pr(Y \geq y)
= F_{XY}(x | y) S_y(y),
\] (7)

where the conditional probabilities are assumed to exist, i.e. \(S_y(y) > 0\).

The input-oriented efficiency score, \(\theta(x, y)\), for \((x, y) \in T\) is defined for \(\forall Y | S_y(y) > 0\) as:

\[
\theta(x, y) = \inf \{\theta | F_{XY}(\theta x | y) > 0\} = \inf \{\theta | H_{XY}(\theta x, y) > 0\}.
\] (8)

In the latter setting, the conditional distribution \(F_{XY}(\cdot | y)\) acts as the feasible set of input values, \(X\), for a DMU exhibiting the output level \(y\). Given the free disposability assumption, the lower boundary of this set (in a radial sense) renders the Farrell-efficient frontier.

A non-parametric estimator is obtained by replacing \(F_{XY}(x | y)\) by its empirical version:

\[
\hat{F}_{XY,k}(x | y) = \sum_{k=1}^{K} I(X_k \leq x, Y_k \geq y) \sum_{k=1}^{K} I(Y_k \geq y),
\] (9)

where \(I(\cdot)\) is the indicator function.

For the output orientation, the Farrel efficiency score is computed as:

\[
\bar{\lambda}(x, y) = \sup \{\lambda | S_{Y\lambda}(\lambda y | x) > 0\} = \sup \{\lambda | H_{XY}(x, \lambda y) > 0\},
\] (10)
with the following empirical estimator of \( S_{yx}(y|x) \):

\[
\hat{S}_{yx}(y|x) = \frac{\sum_{k=1}^{K} I(X_k \leq x, Y_k \geq y)}{\sum_{k=1}^{K} I(X_k \leq x)}.
\]

(11)

Indeed, \( \hat{\lambda}_{FDH}(x,y) = \sup \{ \lambda \mid (x, \lambda y) \in T_{FDH} \} = \hat{\lambda}(x,y) = \sup \{ \lambda \mid \hat{S}_{yx}(\lambda y|x) > 0 \} \).

3. Preliminaries for the estimation of the order–\( \alpha \) frontiers

It is due to Y. Aragon et al. (2005), that the definitions of the efficiency scores given by Eqs. 8 and 10 are based on the order one quantiles of the laws of \( X \) given \( y \leq Y \) and \( Y \) given \( X \leq x \) respectively. Naturally, they proposed a concept of production function of continuous order \( \alpha \in (0,1] \). Note that the concept of the order–\( m \) frontier (Cazals, 2002) is related to the discrete parameter, \( m \). The parameter \((1-\alpha) \times 100\% \) thus indicates the probability that a certain observation is dominated by those producing at least the same amount of outputs (resp. using at most the same amount of inputs) even after the inputs (resp. outputs) are contracted (resp. augmented) with respect to the production frontier. Indeed, the underlying production remains unaltered, whereas the order–\( m \) frontiers are defined in terms of the randomly drawn samples.

A. Daouia and L. Simar (2007), therefore, introduced the order–\( \alpha \) conditional efficiency measures for multi–input and multi–output technology. The \( \alpha \)–quantile input efficiency score for the DMU \((x,y) \in T\) is defined as:

\[
\theta_{\alpha}(x,y) = \inf \{ \theta \mid F_{xy}(\theta x | y) > 1 - \alpha \},
\]

(12)

where \( y \) is such that \( S_{y}(y) > 0 \) and \( \alpha \in (0,1] \). Similarly, the \( \alpha \)–quantile output efficiency score for the DMU \((x,y) \in T\) is defined as:

\[
\lambda_{\alpha}(x,y) = \sup \{ \lambda \mid S_{yx}(\lambda y | x) > 1 - \alpha \},
\]

(13)

where \( x \) is such that \( F_{x}(x) > 0 \) and \( \alpha \in (0,1] \).

The measures described by Eqs. 12–13 can be estimated by plugging in the empirical estimators (cf. Eqs. 9 and 11). Therefore, the estimators of the input and output efficiency scores are

\[
\hat{\theta}_{\alpha,n}(x,y) = \inf \{ \theta \mid \hat{F}_{xy}(\theta x | y) > 1 - \alpha \},
\]

(14)

\[
\hat{\lambda}_{\alpha,n}(x,y) = \sup \{ \lambda \mid \hat{S}_{yx}(\lambda y | x) > 1 - \alpha \}.
\]

(15)

These estimators, in turn, are computed as follows (Daouia, 2007). Let \( M_{y} = \sum_{k=1}^{K} I(Y_k \geq y) > 0 \) and define

\[
\xi_k = \max_{i=1,2,...,p} \left\{ \frac{X_{ik}}{x_i} \right\}, k = 1,2,...,K.
\]

(16)

For \( l=1,2,...,M_{y} \), denote by \( \xi_{(l)}^y \) the permutation of the observations \( \xi_k \) such that \( Y_k \geq y \): \( \xi_{(1)}^y \leq \xi_{(2)}^y \leq \ldots \leq \xi_{(M_y)}^y \). Then we have
\[ \hat{F}_{X|y,k} (\theta | x) = \frac{\sum_{I(y \geq y)} I(X_k \leq \theta x)}{M_y} = \frac{\sum_{I(y \geq y)} I(\xi_k \leq \theta)}{M_y} = \frac{\sum_{l=1}^{M_y} I(\xi_{(l)}^y \leq \theta)}{M_y} \]

(17)

\[ \hat{\theta}_{\alpha,n} (x, y) = \begin{cases} 
\Xi_{(l)}^y & \text{if } (1 - \alpha)M_y \in \mathbb{Z}^* \\
\Xi_{[((1 - \alpha)M_y)+1]} & \text{otherwise}
\end{cases} \]

(18)

Accordingly, \[ \hat{\lambda}_{\alpha,n} (x, y) = \begin{cases} 
\psi_{\max} & \text{if } \alpha N_x \in \mathbb{N}^* \\
\psi_{[\alpha N_x]+1} & \text{otherwise}
\end{cases} \]

(20)

Note that \( \hat{\lambda}_{\alpha,n} \) and \( \hat{\theta}_{\alpha,n} \) converge to the associated FDH estimators as \( \alpha \to 1 \). The FEAR package (Wilson, 2008) was employed to obtain the quantile–based efficiency measures.
4. Data used

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 technical efficiency was assessed in terms of the input and output indicators commonly employed for agricultural productivity analyses, see, for instance, a study by Š. Bojnec and L. Latruffe (2008). 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. The last two variables were deflated by respective real price indices provided by Eurostat. On the other hand, the three output indicators represent crop, livestock, and other outputs in Litas (Lt), respectively. The aforementioned three output indicators were deflated by respective price indices and aggregated into a single one. The aforementioned instance of aggregation was implemented in order to ensure that the randomly drawn output values are reasonable ones across farms of the different specialization.

In order to identify the differences in efficiency across certain farming types, the farms were classified into the three groups in terms of their specialization. Specifically, farms 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 a residual category called mixed farming.

5. Results

The input- and output-oriented measures of the order–α efficiency were implemented to analyse the farm performance with respect to different quantiles. These quantiles, indeed, enable to analyse the variation of the observed data and estimate the level of efficiency. Let \( \hat{q}_i^\alpha \) and \( \hat{q}_o^\alpha \) denote the input and output frontiers (quantiles) of the arbitrary order \( \alpha \), respectively.

The input frontiers were estimated for \( \alpha = \{0.8, 0.85, 0.9, 0.95, 0.99, 0.995, 0.999, 1\} \). Fig. 1 relates the \( \alpha \)-level with the resulting share of observations outside the production frontier. Obviously, the share of farms outside the production frontier did not decrease significantly for \( 0 \leq \alpha \leq 0.95 \). This finding implies that the quantiles, \( \hat{q}_i^\alpha \), associated with the latter level of \( \alpha \) were rather tight and not perturbed by the outliers. In this region the crop farms featured the highest share of the observations inside the production frontier (specifically, 17% at \( \alpha = 0.95 \)), whereas the livestock farms were peculiar with the lowest one (2% at the same \( \alpha \)-level). The quantiles \( \hat{q}_o^\alpha \) with \( \alpha \geq 0.95 \) were influenced by the outliers to a higher extent and thus enveloped higher share of the observations. At \( \alpha = 0.999 \), some 4%, 6%, and 16% of crop, livestock, and mixed farms remained operating outside the production frontier. Therefore, the share of the specialized crop and livestock farms diminished at a faster rate than that of mixed farms.
The output–oriented models yielded a somehow different pattern of quantiles. The quantiles \( \hat{q}_\alpha \) with \( 0 \leq \alpha \leq 0.9 \) were rather compact and not perturbed by the outliers (Fig. 2). At \( \hat{q}_{0.9} \), some 81%, 94%, and 72% of the crop, livestock, and mixed farms were operating outside the production frontier and therefore were considered as super–efficient ones. The share of observations outside the production frontier decayed rapidly for \( \hat{q}_\alpha \) with \( \alpha > 0.9 \). In the output–oriented, framework the mixed farming appeared to feature lowest share of observations outside the quantiles, \( \hat{q}_\alpha \). Specifically, some 11% of the mixed farming observations remained outside the quantile \( \hat{q}_{0.995} \), whereas this percentage dropped to nil for \( \hat{q}_{0.999} \). Anyway, the livestock farming exhibited the highest share of observations outside the production frontier for \( \alpha \leq 0.999 \).
The mean efficiency scores were estimated for each farming types across various $\alpha$–levels. Furthermore, the means were computed for input– and output–oriented models. Figs. 3 and 4 present the results for each farming type.

Fig. 3 clearly indicates that crop farming was generally less efficient if compared to the other farming types for all values of $\alpha$. An average livestock or mixed farm was super–efficient (i. e. the mean efficiency score exceeded unity) for $0.995 \leq \alpha \leq 0.99$, whereas the mean efficiency of crop farms exceeded unity at $0.99 \leq \alpha \leq 0.99$. The mixed farming was the most efficient farming type for $0.9 \leq \alpha \leq 0.9$ and the livestock farming was the most efficient farming type for $\alpha \geq 0.99$. Indeed, the difference between the mean efficiency scores further increased as the values of $\alpha$ approached unity. The crop farming remained the least efficient farming type in terms of the mean efficiency scores for all values of $\alpha$. The FDH estimates of efficiency scores were obtained at $\alpha = 1$. The mean values were 0.8, 0.92, and 0.86 fort crop, livestock, and mixed farms, respectively. These scores can be interpreted as factors of the input contraction required to ensure efficiency, e. g. an average crop farm should contract its inputs by some 20%.
The crop farming was the least efficient faring type for all values of $\alpha$ in the output–oriented model (Fig. 4). Indeed, the output efficiency scores lower than unity indicate super–efficiency, whereas those above unity can be interpreted as the expansion of outputs needed to ensure efficiency. At the other end of spectrum, the livestock farms features mean efficiency scores below unity for $\alpha \leq 0.995$, whereas the other farming types exhibited super–efficiency for $\alpha \leq 0.95$. The quantile $q_{0.999}$ virtually coincided with the FDH frontier. At this $\alpha$–level the output the mean efficiency scores were 1.27, 1.09, and 1.22 for crop, livestock, and mixed farms, respectively. These figures indicate that crop farms should expand their output by 27% on average and so on.

The differences between mean efficiencies associated with different farming types were smaller in the input orientation if compared to those in the output orientation. This finding might indicate that the observations are not distributed evenly inside the production frontier (surface).

6. Conclusions

1. The carried out quantile frontier analysis implied that the share of observations outside the production frontier did not decrease significantly up to $\alpha = 0.8$, i. e. the frontier established after eliminating some 20% of the best performing observations given they produce more outputs or consume less inputs depending on the model’s orientation.

2. Irrespectively of the model’s orientation, the probabilistic analysis of the efficiency of Lithuanian family farms suggested that the crop farming was the least efficient farming type. This result was obtained for all values of $\alpha$.

3. At the other end of spectrum, the mixed and livestock farming featured the highest mean efficiency scores for the input-oriented model. Therefore, these farming types might be considered as similar in terms of input consumption patterns. Note that the mixed farms exhibited the highest efficiency scores for $\alpha \leq 0.9$, whereas the
livestock farms appeared to more efficient as the $\alpha$ increased. Anyway, the livestock farming remained efficient along all values of $\alpha$ for the output-oriented model. These findings implied that the observations are not distributed evenly inside the production frontier (surface).

4. Even though livestock farming appeared to be a relatively efficient farming type, the number of livestock is decreasing Lithuania. Such changes might be fuelled by both economic and social developments in Lithuania. Therefore, the appropriate measures aimed at fostering livestock farming in Lithuania would contribute to increase in agricultural efficiency and productivity.

Acknowledgments

This research was funded by the European Social Fund under the Global Grant measure.

References

LIETUVOS ŪKININKŲ ŪKIŲ VEIKLOS SANTYKINIO EFEKTYVUMO VERTINIMAS
ESANT NEAPIRĖŽTUMUI: ALFA-KVANTILINĖS RIBOS

Tomas Baležentis, Rasa Melnikienė, Mette Asmild,
Jens Leth Hougaard, Aistė Galnaitytė
Lietuvos agrarinės ekonomikos institutas

Įteikta 2013 11 11; priimta 2013 12 10

Santrauka


Raktiniai žodžiai: efektyvumas; ūkininkų ūkiai; alfa kvantilinė riba; veiklos analizė.

JEL kodai: C14, C44, D24, Q12.