MODELING MILK COST IN ESTONIA: A STOCHASTIC FRONTIER ANALYSIS APPROACH

R. Põldaru, J. Roots

Estonian University of Life Sciences

ABSTRACT. *This paper presents a formulation of stochastic frontier models for milk cost in Estonia. Two distinct models of milk cost were investigated. A balanced panel of 45 Estonian farmers observed during the period 2001 to 2006 was used. For the models parameter estimation a computer program FRONTIER Version 4.1 was used. The results for various specifications were compared and discussed. The results from stochastic frontiers model analysis were compared with the results of OLS. Predicted cost efficiencies of the Estonian farmers were compared under the different model specifications. This analysis demonstrated that stochastic frontier analysis (SFA) can be implemented for parameter estimation of econometric models and for predicting the cost efficiency of milk cost in Estonian farms.*

Keywords: *stochastic frontier analysis, cost (economical) efficiency, panel data, milk cost*

Introduction

Estonia is one of the new members of the European Union. The EU enlargement means for East European countries a lot of changes in their agriculture. These changes are at the political, economical and technical level. This means that information systems on agriculture (databases, models etc) have to move along with those changes. Consequently, the economic models in Estonia have either to be created, developed or renewed, and must be harmonised with the European requirements. Hopefully, we can use new information technology to lead such evolutions.

We recognise that there is a variation in the behavioural characteristics of the agricultural production systems over time as well as between countries. The diverse nature of agricultural production systems and agri-food markets across the EU poses a challenge to anyone seeking to develop a model that can be used to analyse policy at EU and member state level.

Improving the competitiveness of Estonian agriculture is the priority objective of agricultural policy. The outcome and impacts of those policy actions will strongly depend on developments of the agricultural world markets. The dairy sector is the most competitive commodity of Estonian agriculture. Consequently, the need to make Estonian dairy farms more competitive is obvious.

At the Estonian University of Life Sciences (Institute of Economics and Social Sciences), we have investigated the possibilities of some new Data Mining (DM) methods and have some experience in implementing algorithms used in DM packages. We have used various methods for estimating the parameters of econometric model of grain yield and milk cost.

We have used Bayesian statistical methods in Põldaru and Roots (2001b), neural networks in Põldaru and Roots (2003), principal components method in Põldaru and Roots (2001a), decision trees and rules – CART – (Classification And Regression Trees) in Põldaru *et al.* (2003b), association rules discovery in Põldaru *et al.* (2003a), fuzzy regression in Põldaru *et al.* (2004a); and support vector machines regression (Põldaru *et al.*, 2004c; Põldaru *et al.*, 2004d; Põldaru *et al.*, 2005) for estimating the parameters of econometric model of grain yield and milk cost.

Recently Journal of Productivity Analysis published special issue discussing the productivity and efficiency problems of countries that might be moving from command economies to market economies (Li *et al.*, 2008).

In recent decades, the interest of econometricians for new models and methods has increased substantially, including the stochastic frontier analysis (SFA).

The stochastic frontier analysis (SFA) was previously used to model agricultural production (Coelli, Battese, 1996; Hadri, Whittaker, 1999), grain production (Battese, Broca, 1997; Põldaru, Roots, 2004b), milk production (Reinhard, *et al.*, 2000; Lawson *et al.*, 2004; Abduali, Tietje, 2007), meat production (Sharma *et al.*, 1997) and wool production (Fraser, Horrace, 2003). Recently an extensive overview of empirical studies of technical efficiency in farming was published (Bravo-Ureta *et al.*, 2007).

In this paper we consider SFA as a method for econometric model parameter estimation and as an instrument to predict economical efficiencies of milk production in Estonian farms. Next we investigate the possible use of two distinct models of milk cost. The study differs from previous studies because it discusses the efficiency of milk production in the country that is moving from command economies to market economies.

In this study we use the approach that is generally preferred in efficiency analyses of agricultural performance, where data noise might be a significant issue (Coelli, 1995).

Two specific econometric models were specified.

The first model (MI) is a neutral stochastic frontier model where farm specific inefficiency explanatory variables are assumed to be independent of the input variables in the production function. The first model is relatively correct (almost all essential independent variables are included in the model, the coefficient of determination, R^2, is high, almost all parameter estimates are significant and acceptable from economic point of view). The parameters of the first econometric model

(MI) were previously estimated implementing ordinary least square regression (OLS) method.

The second model (MII) is a modified stochastic frontier model where farm specific inefficiency explanatory variables are assumed to account for cost inefficiency in production, independent of the input variables in the production function.

For the both stochastic frontier models two alternatives are considered: a) analysis of cross-section data and b) analysis of panel data. The model parameters for different variants of independent variables specification were estimated. For the model parameter estimation, a computer program FRONTIER Version 4.1 was used (Coelli, 1996).

The results for various specifications were compared and discussed. The results from stochastic frontiers model analysis were compared with results of previous analyses.

The data is a balanced panel of 45 Estonian milk producers drawn from FADN (Farm Accountancy Data Network) observed during the period 2001 to 2006.

This paper is organized as follows. The next section describes the frontier cost models used. Section 2 describes the data for the empirical analyses. Section 3 presents and discusses the results. Section 4 summarizes and gives conclusions.

Stochastic frontier models for milk cost in Estonian farms

In this paper the standard stochastic frontier cost function models (M I) for panel (or cross-sectional) data was used. This first model is described more thoroughly in Battese and Coelli (1992). The model may be expressed as:

$$Y_{it} = \beta_0 + \sum_{j=1}^{K} \beta_j \cdot x_{jit} + \left(V_{it} + U_{it}\right) \qquad (1)$$

where:

Y_{it} is the milk cost of the i-th farm in the t-th time period;

x_{jit} is the j-th input quantity of the i-th farm in the t-th time period;

β is $K \times 1$ vector of unknown parameters;

the V_{it} are random variables which are assumed to be independent identically distributed normal random variables (iid $N\left(0, \sigma_V^2\right)$), and independent of the U_{it};

$$U_{it} = \left(U_i \cdot \exp\left(-\eta \cdot (t - T)\right)\right) \qquad (2)$$

the U_i are non-negative random variables which are assumed to account cost of inefficiency in milk cost model and are assumed to be iid $N\left(\mu, \sigma_U^2\right)$;

η is a parameter to be estimated using panel data.

The parameterization of Battese and Corra (1977) who replaced σ_V^2 and σ_U^2 with

$\sigma^2 = \sigma_V^2 + \sigma_U^2$ and $\gamma = \sigma_U^2 / \left(\sigma_V^2 + \sigma_U^2\right)$ is followed. (2a).

The cost (economical) efficiency of a given farm (at a given time period) is defined by Battese and Coelli (1992) as the ratio of its mean cost (milk cost) to the corresponding mean cost if the farm utilized its levels of inputs most efficiently (as the ratio of minimum feasible cost to observed expenditure). The measure of cost (economical) efficiency relative to the cost frontier (1) is defined as:

$$CE_i = E\left(Y_i^* / U_i, X_i\right) / E\left(Y_i^* / U_i = 0, X_i\right) \qquad (3)$$

where,

Y_i^* is the cost (milk cost) of the i-th farm, which will be equal to Y_i when the dependent variable is in original units and will be equal to $\exp(Y_i)$ when the dependent variable is in logs. CE_i will take a value between one and infinity in the cost function case.

The imposition of one or more restrictions upon this model formulation can provide a number of the special cases of this particular model which have appeared in the literature. Setting η to be zero provides the time-invariant model (variant MI-2). Furthermore, restricting the formulation to a full (balanced) panel of data gives the production function assumed in Battese and Coelli (1988). The additional restriction of μ equal to zero reduces the model to variant MI-1. The restriction of T=1 return to the original cross-sectional (variant MI-1), half-normal formulation of Aigner, Lovell and Schmidt (1977). Obviously a large number of permutations exist. For example, if all these restrictions excepting μ=0 are imposed, the model suggested by Stevenson (1980) results. Furthermore, if the cost function option is selected, we can estimate the model specification in Hughes (1988) and Schmidt and Lovell (1979) specification, which assumed allocative efficiency. These latter two specifications are the cost function analogues of the production functions in Battese and Coelli (1988).

There are obviously a large number of model choices that could be considered for any particular application. For example, does one assume a half-normal distribution (variant MI-4) for the inefficiency effects or the more general truncated normal distribution (variant MI-3). If panel data are available, should one assume time-invariant or time-varying efficiencies? If such decisions must be made, it is recommended that a number of the alternative models be estimated and that a preferred model be selected using likelihood ratio tests.

The second model (MII) is described more thoroughly in Battese and Coelli (1992). The model may be expressed as:

$$Y_{it} = \beta_0 + \sum_{j=1}^{K} \beta_j \cdot x_{jit} + \left(V_{it} + U_{it}\right) \qquad (4)$$

where Y_{it}, x_{it}, and β are as defined earlier;

the V_{it} are random variables which are assumed to be iid $N(0, \sigma_V^2)$, and independent of the

U_{it} which are non-negative random variables which are assumed to account for cost inefficiency in production and are assumed to be independently distributed as truncations at zero of the $N(m_{it}, \sigma_U^2)$ distribution; where:

$$m_{it} = \delta \cdot z_{it} \quad \text{or} \quad m_{it} = \delta_0 + \sum_{j=1}^{P} \delta_j \cdot z_{itj} \qquad (5)$$

where z_{itj} is the j-th input quantity of the i-th farm in the t-th time period (variables which may influence the efficiency of a firm); and

δ is an $1 \times P$ vector of parameters to be estimated.

This model specification also encompasses a number of other model specifications as special cases. If we set $T=1$ and z_{it} contains the value one and no other variables (i.e. only a constant term), then the model reduces to the truncated normal specification, where δ_0 (the only element in δ) will have the same interpretation as the μ parameter in first model MI (variant MI-2). It should be noted, however, that the model defined by (4) and (5) does not have the model defined by (1) as a special case, and neither does the converse apply. Thus these two model specifications are non-nested and hence no set of restrictions can be defined to permit a test of one specification versus the other.

Data

In this study we utilize data describing the production activities of 45 highly specialised dairy farms (Decision Making Units – DMU) that were in the Estonian Farm Accountancy Data Network (FADN) all of the 2001–2006 period. The FADN is a stratified random sample. Stratification is based on economic farm size, age of the farmer, region, and type of farming. We have a total of 270 observations in this balanced panel, and so each farm appears 6 times in the panel. The period 2001 is chosen because detailed information at each farm is available from 2001 onwards.

A panel contains more information than does a single cross section. Consequently it is to be expected that access to panel data will either enable some of the strong distributional assumptions used with cross-sectional data to be relaxed or result in estimates of cost efficiency with more desirable statistical properties. The fundamental problem is that in a single cross section we get to observe each producer only once, and this severely limits the confidence in our cost efficiency estimates.

In the selection of independent variables we must address the trade-off between using technical details by applying more inputs and adding the risk of multicollinearity on the one hand, and diminishing the inputs and sacrificing potentially useful information on the other hand.

Note that we use the preliminary analysis to select variables that have a significant influence on milk cost function.

The linear cost function defined in equations (1) and (4) is estimated using ordinary least squares regression. At this step, we identify inputs expected to have a significant influence on cost function before the frontier production function is estimated using maximum likelihood estimation.

The dependent variable is a milk cost per kg of milk output (Y), and independent variables are average milk yield per cow (x_1), labour input per 100 kg milk (x_2), total feed cost per 1 kg of milk (x_3) and number of milking cows in herd (x_4).

The characteristics of the selected data are summarised in Table 1.

Table 1. Data summary statistics

Definitions of variables	Measure	Characteristics	Years						All panel
			2001	2002	2003	2004	2005	2006	
Y-milk cost		Mean	2.90	3.82	3.70	4.38	4.70	4.62	4.02
per kg of milk	kroons	St.dev.	0.77	1.24	1.01	1.06	1.16	1.09	1.22
output		Minimum	1.44	2.36	2.18	2.75	3.23	2.88	1.44
		Maximum	4.46	8.39	6.38	6.79	7.47	8.14	8.39
		Mean	5530	5600	5597	5887	6225	6516	5892
x_1 – average milk yield		St.dev.	1401	1545	1587	1473	1518	1522	1288
per cow	kg	Minimum	2654	2363	2228	2825	2502	3267	2228
		Maximum	8327	8475	8508	8549	9417	9735	9735
		Mean	4.19	4.00	3.60	3.24	3.03	2.68	3.46
x_2- labour input per		St.dev.	1.74	1.83	1.68	1.46	1.41	1.20	1.62
100 kg milk	hours	Minimum	1.66	1.50	1.10	1.07	1.05	0.80	0.80
		Maximum	9.39	10.22	9.40	8.26	7.44	5.62	10.22
		Mean	1.15	2.12	1.91	2.31	2.35	2.30	2.02
x_3- total feed cost per		St.dev.	0.40	0.84	0.66	0.72	0.75	0.91	0.79
1 kg of milk	kroons	Minimum	0.60	0.65	0.49	1.22	1.28	1.17	0.49
		Maximum	2.39	5.73	3.58	3.80	3.89	5.04	5.73
		Mean	114	118	120	123	128	134	123
		St.dev.	147	154	150	153	161	169	155
x_4- number of milking		Minimum	20	22	25	25	27	25	20
cows in herd		Maximum	663	731	710	702	779	821	821

One feature of the sample is that the mean value of dependent variable (milk cost per kg of milk output) is changing. In the years 2001–2005 average milk cost with fluctuations increased from 2.90 kroons to 4.70 kroons per kg and then decreased to 4.62 kroons per kg in 2006. Nearly analogously changes the mean value of independent variable x_3 – total feed cost per 1 kg milk. At the same time mean values of other independent variables are changing with almost constant trend. The average milk yield per cow (x_1) increased from 5,530 kg per cow in 2001 to 6,516 kg in 2006. The total labour input x_2 (hours per 100 kg milk) has decreased essentially (from 4.19 hours in 2001 to 2.68 hours in 2006). Consequently, the labour input decreased 1.6 times. The average number of milking cows in herd, x_4, increased from 114 cows in 2001 to 134 cows in 2006. This increase is moderate as compared to decrease of labor input.

Consequently, the milk production in Estonian farms is not obtained a stable state. Table 1 show, that the most critical is the year 2004.

This situation may be explained by the features of moving from the socialist economic system to market economies:

- The process of moving from command economies to market economies in Estonia is not ended yet.
- Beginning from 2004 Estonia is one of the new members of the European Union.
- Before 2004 the prices (including prices of inputs for milk production) increased.
- At same time the efficiency of milk production is rising (labour use decreases).
- After 2004 the economic situation changed.

Considering circumstances described before, one additional independent variables was included in the milk cost model: the trend variable – x_5.

Results and discussion

The frontier functions (1) and (4) are estimated for several alternative models. To derive our preferred functional form we estimated six specifications (four alternatives for MI and two alternatives for MII).

The specifications for MI alternatives are presented in Table 2.

Table 2. Description of model MI specifications

Definition of specification	Description of the specifications	Parameters mu μ	eta η	T
MI-1	cross-sectional, half-normal inefficiency, time-invariant model	0	0	1
MI-2	cross-sectional, truncated-normal inefficiency, time-invariant model	y	0	1
MI-3	panel-data, truncated-normal efficiency	y	y	6
MI-4	panel-data, half-normal efficiency, time-invariant model	0	0	6

The specifications for MII alternatives are presented in Table 3.

Table 3. Description of model MII specifications

Definition of specification	Description of the specification	Parameters mu μ	eta η	T	del0 δ_0	del1 δ_1	del2 δ_2
MII-1	panel-data, truncated-normal	y	0	6	y	y	y
MII-2	panel-data, truncated-normal	y	0	6	y	y	0

The specifications and maximum-likelihood estimates of the parameters in the milk cost stochastic frontier function defined by equations (1) and (4) for alternative specifications are given in Table 4.

Table 4. Maximum-likelihood estimates for parameters of stochastic frontier production function of milk costs for different model specifications

Variable	Parameter	OLS	MI-1	MI-2	MI-3	MI-4	MII-1	MII-2
Intercept	β_0	0.440	0.074	0.387	0.067	-0.182	0.620	0.537
Milk yield	β_1	0.000028	–0.000006	–0.000024	–0.000022	–0.000016	–0.000012	–0.000002
Labour input	β_2	0.143	0.140	0.135	0.120	0.106	0.127	0.128
Feed cost	β_3	1.062	0.925	0.902	1.046	1.045	1.024	1.022
Number of cows	β_4	0.0015	0.0013	0.0013	0.0011	0.0010	0.0013	0.0013
Trend	β_5	0.168	0.181	0.184	0.116	0.170	0.000	0.000
Sigma-squared	σ^2	0.446	1.135	2.771	1.141	1.012	0.937	0.908
Gamma	γ		0.958	0.969	0.886	0.866	0.879	0.883
Mu	μ		0	-2.730	0.319	0	0.0	0.104
Eta	η		0	0	-0.076	0	0	0
Inefficiency								
Average	$\overline{u}$		0.829	0.680	0.789	0.753	0.795	0.818
Minimum	u_{min}		0.091	0.113	0.038	0.039	0.119	0.116
Maximum	u_{max}		3.027	2.868	3.228	2.941	3.051	3.064
R^2		0.707	0.586	0.571	0.612	0.668	0.500	0.497

Table 4 also reports the parameters OLS estimates for alternative MI and presents result summaries of the results of various MI and MII alternatives. Summary characteristics for various alternatives are: sigma-squared – σ^2, gamma – γ, mu – μ, eta – η, and summary characteristics of cost (economical) economical inefficiency (minimum, mean and maximum) and coefficient of determination – R^2.

The coefficients of the exploratory variables β_i in the milk cost model (the stochastic frontier cost function) are of particular interest to this analysis.

Next we analyse the parameter estimates in Table 4.

The results in Table 4 indicate that the OLS estimates (OLS) and SFA estimates (MI-1 and MI-2) are similar, whereas the estimates for feed cost (β_3) are essentially equivalent. The estimates for intercept (β_0), labour input (β_2), number of cows (β_4) do not differ essentially.

Comparing the signs of parameter estimates for different mode specifications, one should conclude that only once the sign is changing. In the case of independent variable – milk yield per cow (β_1). The estimates of other independent variables for all variants have the same sign – positive or negative. Consequently the SFA models are robust.

It is important to note that the estimate sign for milk yield per cow (β_1) is positive for OLS, and sign is negative for all SFA alternatives. The economic theory and practice assert that the model parameter should be negative for the variable – milk yield per cow. Consequently, in the case of OLS the estimate for independent variable, milk yield per cow, is not adequate and SFA are preferred.

Comparing the parameter estimates for models MI and MII, one should conclude that parameter estimates practically do not differ. The models parameters signs for all specifications (variants) are the same.

Finally it may be concluded that the OLS and SFA estimates don't differ significantly.

Next we analyse the summary characteristics for SFA models in Table 4.

First, we analyse the characteristics for model MI (see equation (1) and Table 2). Comparing the summary characteristics in Table 4 it may be concluded that the characteristics differ in different cross-sectional data (MI-1 and MI-2) and panel data (MI-3 and MI-4) models. Comparing the estimates of sigma squared, σ^2 (σ^2 is calculated using equation (2a)), for model MI variants, one should conclude that estimates differ. In the case of alternative MI-4 (panel data, half-normal inefficiency distribution, time-invariant model) the value of σ^2 is minimal (σ^2=1.012) and in the case of alternative MI-2 (cross-sectional data, truncated-normal inefficiency distribution, time-invariant model) the value of σ^2 is maximal (σ^2=2.771). The values of σ^2 in Table 4 differ approximately 2.7 times. Comparing the estimates of sigma squared, σ^2 ,for model MII variants, one should conclude that estimates do not differ.

The values of parameter gamma, (γ), do not differ substantially. The value 0.969 of the parameter gamma, (γ), in estimated model MI-2 is maximal. It implies that the predicted variance of inefficiency (see equation (2a)) is estimated to have a value higher approximately by a factor of 30 than the estimated value of variance of random variable V. That difference is essential. The value 0.883 of the parameter gamma, (γ), in estimated model MII-2 is minimal. It implies that the predicted variance of inefficiency is estimated to have a value higher approximately by a factor of 7.5 than the estimated value of variance of random variable V. Consequently, in different variants the inefficiency component involve different amount of information.

Because the estimates for the parameter, η, is negative, the inefficiency of milk cost for Estonian farmers tend to decrease over time, according to alternative MI-3.

The values of the coefficient of determination, R^2, are relatively high. The minimal value (0.479) and maximal value (0.668) of R^2 in Table 4 does differ. Thereby in cases of alternatives MI-4 (R^2=0.668) and MI-3 (R^2=0.612) the values of R^2 are higher than in cases of alternatives MII-2 (R^2=0.497) and MII-1 (R^2=0.500). But at the same time the values of R^2 for SFA models are lower than for OLS model (R^2=0.707).

Next we analyse the cost inefficiency characteristics in Table 4. It should to be mentioned, that cost inefficiency is measured in units of dependent variable (in units of milk cost). Consequently inefficiency in Table 4 is measured in kroons per kg of milk output. The average value of inefficiency, $\overline{u}$, for different model variants does not differ substantially. The average inefficiency ranges between 0.680 (MI-1) and 0.829 (MI-2) for cross-section models, and ranges between 0.753 (MI-4) and 0.818 (MII-2) for panel data models. Consequently, Estonian farmers on an average have a reserve to reduce milk cost approximately by 80 cents.

It should to be noted, that inefficiency is producer (farmer) specific characteristic. Comparing the inefficiency variability characteristics (minimum and maximum) in Table 4 it may be concluded that the characteristics differ in different cross-sectional data and panel data models. In the case of MI the inefficiency ranges between 0.113 and 2.868 in alternative MI-2 and, between 0.038 and 3.228 in alternative MI-3. The predicted inefficiencies for model MI-2 exhibit less variability than in MI-3. In the case of MI-3 minimal inefficiency equals 0.038 (a reserve to reduce milk cost is only by 4 cents) and the maximal inefficiency equals 3.228 (a reserve to reduce milk cost is very large – by 3.23 kroons). The last value is authentic (is not astonishing), while the two maximal values of milk cost in Table 1 are equal to 8.14 and 8.39. Specific analysis shows that in present case the actual milk cost is equal to 7.26. It should to be noted, that the same farm is most ineffective in all model variants. Consequently, there are reserves. In the case of most effective farm (U=0.038) the actual milk cost is equal to 1.96 kroons per kg of milk output. So low was milk cost in year 2001.

For the variants MII-1 and MII-2 the variability of inefficiency is practically the same.

Next we analyse the distributions of inefficiency for considered alternative models. The Figure 1 shows the histograms of cost inefficiency for different model variants.

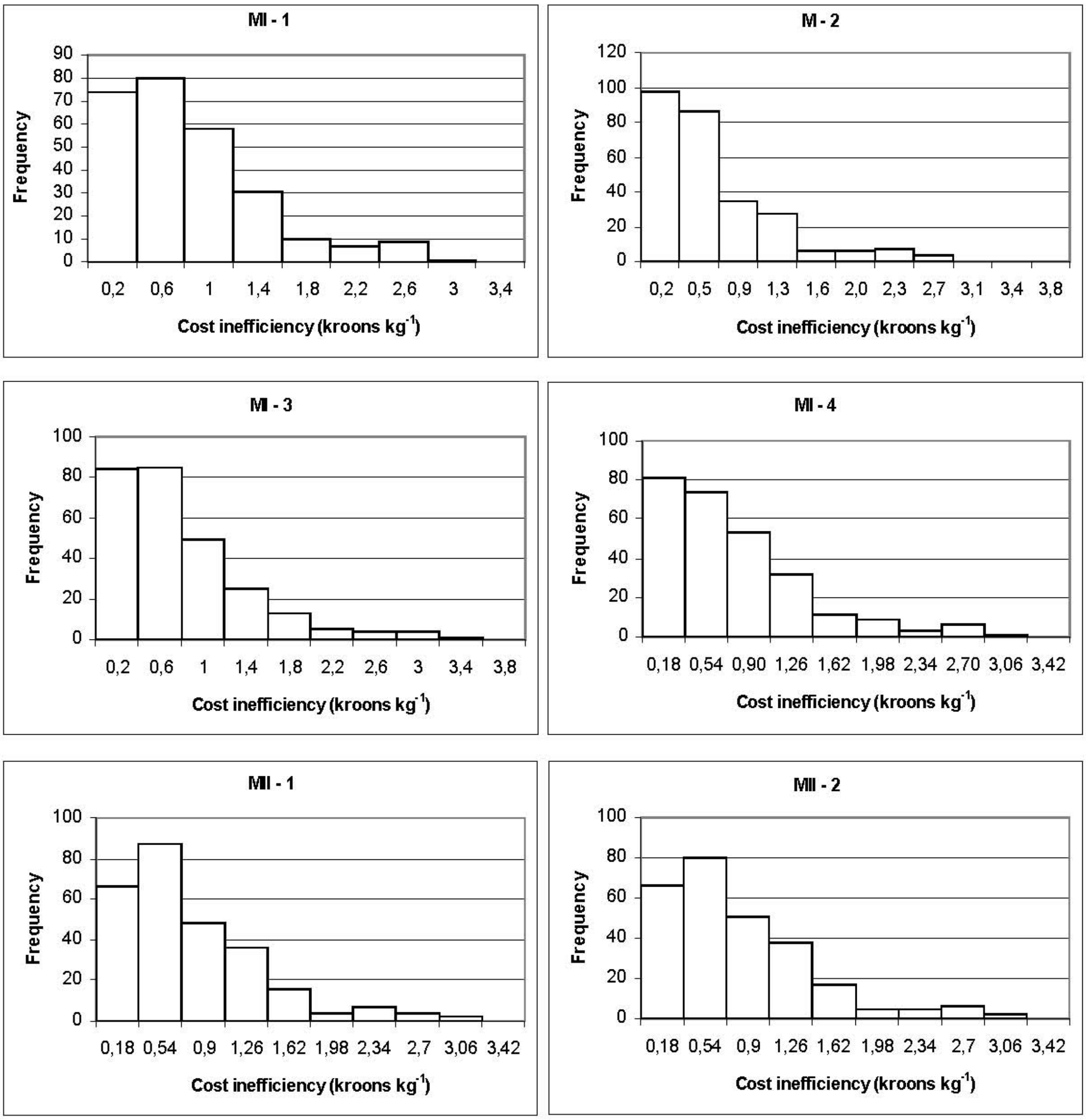

Figure 1. Histograms of cost inefficiency for different model variants

As indicated in Table 4, in three variants (MI-1, MI-4 and MII-1) half-normal distribution for random variable U_i (inefficiency) was assumed (parameter μ=0). As seen from Figure 1, histograms for those variants have asymmetric character with maximum at zero. Consequently, in those cases the estimates of inefficiency have half-normal distribution. For other variants (MI-2, MI-3 and MII-2) the truncated normal distribution was assumed. Figure 1 shows, that histograms for those variants have asymmetric character with maximum approximately at estimated value of parameter μ (see Table 4). For example, for variant MI-3 the estimated μ=0.319 and from the histogram we can find approximately the same maximum value.

Next we compare estimates of economical efficiency (cost inefficiency) for considered alternative models. For that purpose we check the robustness of our cost (economical) efficiency results. A simple test of whether the rank of farms (DMU-s) is robust to different model specification is to estimate the Spearman Rank Correlation coefficient between the various model alternatives (Frazer and Horrace 2003).

We estimated cost (economical) efficiency for all the model alternatives and derived the rank of the farms (DMU-s). Then, rank correlation coefficient was estimated for all pairs of model alternatives; the results are reported in Table 5.

Table 5 provides also rank correlation coefficients between ranks of farms (DMU-s) for different model alternatives and ranks of farms obtained using OLS model. For the OLS model the rank of the farms (DMU-s) was derived using regression residuals. The model alternatives in Table 4 are grouped. The two first models (MI-1 and MI-2) use cross-sectional data and four last models (MI-3, MI-4, MII-1 and MII-2) panel data.

Table 5. Spearman rank correlation coefficients

Model variant	OLS	Model variants					
		MI-1	MI-2	MI-3	MI-4	MII-1	MII-2
OLS	1.000						
MI-1	0.983	1.000					
MI-2	0.976	0.999	1.000				
MI-3	0.708	0.744	0.747	1.000			
MI-4	0.742	0.775	0.777	0.988	1.000		
MII-1	0.781	0.813	0.817	0.802	0.762	1.000	
MII-2	0.792	0.823	0.826	0.806	0.768	0.999	1.000

As we can see from the estimates in Table 5, there is a very strong positive relationship across the variants of models estimated. So in the case of using cross-sectional data, there is practically functional relationship across models (MI-1 and MI-2). In the case of using panel data, there is very strong relationship across models (MI-3, MI-4) and (MII-1, MII-2). Thus, despite a difference between the different model specifications, we are able to assume that the order (rank) of efficient/inefficient DMU-s tend to be the same across model alternatives.

Comparing the rank correlation coefficients between ranks of farms (DMU-s) for different model alternatives and ranks of farms obtained using OLS model in Table 5, it may be concluded that the correlation coefficient differ in different cross-sectional data and panel data models. So in the case of using cross-sectional data, there is very strong relationship across models (MI-1 and MI-2) and OLS model, but in the case of panel data models there is a relatively dense relationship across models (MI-3 and MI-4) and OLS model. In the case of model MII there is also a relatively dense relationship across variants (MII-1 and MII-2) and OLS model.

Hence, cost efficiency rankings are fairly robust to model specification for this particular data set. These results are consistent with the existing findings in the frontier literature (Kumbhakar and Lowell (2000) and Frazer and Horrace (2003)).

Conclusions

In this paper we have estimated the stochastic frontier cost function for a panel of milk cost data in Estonian farms and have estimated the cost (economical) efficiency of milk production in Estonian farms.

By comparing the OLS, MI and MII models we may deduce:

In the case of model MI (the inefficiency distributions mean value is constant), the OLS and SFA parameter estimates do not differ significantly; the coefficient of determination, R^2, for SFA models is lower than for OLS models. For the cross-sectional data models the efficiency scores are relatively high (inefficiency scores are relatively low) in the variant MI-2. The rank correlation coefficients for all pairs of model alternatives are very strong. For the panel data models MI-3 and MI-4 the predicted efficiency scores exhibit practically the same variability as in cross-sectional data models and tend to decrease over time. In the case of panel data the SFA models the analysis gives some new information.

In the case of model MII (the inefficiency distributions mean value is different), the OLS and SFA estimates also do not differ significantly. The coefficient of determination, R^2, for SFA models is lower than for OLS model. The efficiency scores exhibit practically the same variability than in the model MI.

This analysis has demonstrated that SFA can be implemented for parameter estimation of econometric models and predicted efficiency scores give new additional information about milk production in Estonian farms.

This analysis showed that Estonian farmers on an average have a reserve to reduce milk cost per kg of milk output approximately by 80 cents.

Regardless of functional form used, the efficiency information that emerges from the analysis is limited to producer-specific estimates of the cost of inefficiency. With a single-equation model, and without input quantity or input cost share data, it is not possible to decompose these estimates into estimates of the cost of input oriented technical inefficiency and the cost of input allocative inefficiency. A decomposition requires additional data and a simultaneous equation model.

Consequently, single-equation cost frontier models are easy to estimate, but they generate limited information. If all that is desired is producer-specific estimates of cost efficiency, single-equation models are adequate for the task.

References

Abdulai, A., Tietje, H. 2007. Estimating technical efficiency under unobserved heterogeneity with stochastic frontier models: application to northern German dairy farms. – European Review of Agricultural Economics, 34, p. 393–416.

Aigner, D.J., Lovell, C.A.K., Schmidt, P., 1977. Formulation and Estimation of Stochastic Frontier Production Function Models. – Journal of Econometrics, 6, p. 21–37.

Battese, D. E., Corra, G.S., 1977. Estimation of Production Frontier Model: With Application to the Pastoral Zone of Eastern Australia. – Australian Journal of Agricultural Economics, 21, p. 169–179.

Battese, G.E., Coelli, T.J., 1988. Prediction of Firm-Level Technical Efficiencies With a Generalised Frontier Production Function and Panel Data. – Journal of Econometrics, 38, p. 387–399.

Battese, D. E., Coelli, T.S., 1992. Frontier Production Functions, Technical Efficiency and Panel Data: with Application to Paddy Farmers of India. – Journal of Productivity Analysis 3, p. 153–169.

Battese, D.E., Broca, S.S., 1997. Functional Forms of Stochastic Frontier Production Functions and Models for Technical Inefficiency Effects: A Comparative Study for Wheet Farmers in Pakistan. – Journal of Productivity Analysis, 8, p. 395–414.

Bravo-Ureta, B.E., Solis, D., Lo´pez, V.H.M., Maripani, J.F., Thiam, A., Rivas, T., 2007. Technical efficiency in farming: a meta-regression analysis. – Journal of Productivity Analysis, 27, p. 57–72.

Coelli, T. 1995. Recent developments in frontier modelling and efficiency measurement. Aust. J. Agric. Econ. 39, 219–245.

Coelli, T., 1996. A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation. – CEPA Working Paper 96/07,
http://www.une.edu.au/econometrics/cepa.htm.

Coelli, T., Battese, G.E., 1996. Identification of factors which influence the technical inefficiency of Indian farmers. – Australian Journal of Agricultural Economics, 40, p. 103–128.

Frazer, I.M., Horrace, W.C., 2003. Technical Efficiency of Australian Wool Production: Point and Confidence Interval Estimates. – Journal of Productivity Analysis 20, p. 169–190.

Hadri, K., Whittaker J., 1999. Efficiency, environmental contaminants and farm size: testing for links using stochastic production frontiers. – Journal of Applied Economics, 11, p. 337–356.

Hughes, M.D., 1988. A Stochastic Frontier Cost Function for Residential Child Care Provision. – Journal of Applied Econometrics, 3, p. 203–214.

Kumbhakar, S.C., Lowell, C.A.K., 2000. Stochastic Frontier Analysis. Cambridge University Press, Cambridge, 333 p.

Lawson, L.G., Agger, J.F., Lund, M., Coelli, T., 2004. Lameness, metabolic and digestive disorders, and technical efficiency in Danish dairy herds: a stochastic frontier production function approach. – Livestock Production Science, 91, p. 157–172.

Li, S., Simar, L., Wilson, P.W., Zelenyuk, W., 2008. Introduction for Journal of Productivity Analysis special issue on transitioning economics. – Journal of Productivity Analysis, 29, p. 77.

Põldaru, R., Roots, J., 2001a. On the Implementation of the Principal Component Regression for the Estimation of the Econometric Model of Grain Yield in Estonian Counties. – Problems and Solutions for Rural Development, International Scientific Conference Reports (Proceedings) (ed Kozlinskis, V.). Latvia University of Agriculture. Jelgava*,* p. 340–345.

Põldaru, R., Roots, J., 2001b. Bayesian Statistics BUGS in the Estimation of the Econometric Model of Grain Yield in Estonian Counties. – Agriculture in Globalising World, Proceedings volume II of International Scientific Conference (ed. J. Kivistik). EAA No. 15/2001,Tartu, p. 178–189.

Põldaru, R., Roots, J., 2003. The Estimation of the Econometric Model of Grain Yield in Estonian Counties Using Neural Networks. – VAGOS, No. 57, Akademija, Kaunas, p. 124–130.

Põldaru R., Roots J., Ruus R., 2003a. A Perspective of Using Data Mining Association Rules in Rural Areas. –, Transactions of the Estonian Agricultural University, No 218, Perspectives of the Baltic States' Agriculture under the CAP Reform, 19–20 September, 2003. Proceedings of International Scientific Conference (ed Kerner, Ü.), Tartu, p. 184–199.

Põldaru R., Roots J. Ruus R., 2003b. A Perspective of Using Data Mining in Rural Areas. – VAGOS, No. 61, Akademija, Kaunas , p. 133–141.

Põldaru R., Roots J. Ruus R., 2004a. Using Fuzzy Regression in Rural Areas. – Economic Science for Rural Development – Possibilities of Increasing Competitiveness, Proceedings of the International Scientific Conference No 7 (eds. Strikis, V., Zandere, E.). Jelgava, p. 43–48.

Põldaru R., Roots J., 2004b. Modelling Grain Yield in Estonian Counties: A Stochastic Frontier Analysis Approach. – *Data Envelopment Analysis and Performance Management, Proceedings of the 4th International Symposium of DEA 5th–6th September 2004 Aston Business School Aston University UK*, p. 261–267.

Põldaru R., Jakobson R., Roosmaa T., Roots J., Ruus R., Viira A-H., 2004c. Support Vector Machine Regression in Estimating Econometric Model Parameters. – *Information Technologies and Telecommunication for Rural Development, Proceeding of the International Scientific Conference Jelgava, Latvia, 6–7 May, 2004. Jelgava,* p. 66–77.

Põldaru R., Roots J., Viira A.-H., 2004d. The estimation of the econometric model of milk yield per cow: a support vector machine regression approach. – *Operations Research 2004 International Conference, Tilburg University, Netherlands, 1–3 September, Tilburg,* p 166–167.

Põldaru R., Roots J., Viira A.-H., 2005. Estimating econometric model of average total milk cost: a support vector machine regression approach. – Economics and Rural Development, 1, p. 23–31.

Reinhard S, Lovell C, Thijssen G., 2000. Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. – European Journal of Operational Research, 121, p. 287–303.

Schmidt, P., Lovell, C.A.K., 1979. Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers. – Journal of Econometrics, 9, p. 343–366.

Sharma, K.R., Leung, P.S., Zaleski, H.M., 1997 Productive efficiency of the swine industry in Hawaii: stochastic frontier vs. Data envelopment analysis. – Journal of Productivity Analysis, 8, p. 447–459.

Stevenson, R.E., 1980. Likelihood Functions for Generalised Stochastic Frontier Estimation. – Journal of Econometrics, 13, p. 57–66.

Piima tootmiskulude (omahinna) modelleerimine Eestis: stohhastilise piiranalüüsi meetodil

R. Põldaru, J. Roots

Kokkuvõte

Antud artiklis käsitletakse Eesti ettevõtete piima arvutuslike tootmiskulude (oma)hinna mudeli koostamist stohhastilise piiranalüüsi meetodil. Artiklis antakse lühiülevaade stohhastilise piiraanalüüsi meetodi olemusest ja kasutamisvõimalustest tootmiskulude modelleerimisel. Töös käsitletakse kahte erinevat piima arvutuslike tootmiskulude (omahinna) mudelit. Mudelite parameetrite hindamiseks on koostatud vastav andmestik, mis kujutab endast tasakaalustatud andmepaneeli 45 ettevõtja kuue aasta (2001–2006) andmetest. Andmestiku koostamisel on kasutatud põllumajandusliku raamatupidamise andmebaasi (FADN) andmeid. Mudeli parameetrite leidmiseks (hindamiseks) kasutati tarkvarapaketti *FRONTIER* (versioon 4.1). Kahe erineva mudeli jaoks koostati kokku 6 erinevat varianti. Artiklis võrreldakse ja analüüsitakse erinevate variantide parameetreid ja efektiivsusnäitajaid. Stohhastilise piiranalüüsi meetodi abil leitud mudeli parameetrid on võrreldus klassikalise regressioonanalüüsi tulemustega. Analüüs näitas, et stohhastilise piiranalüüsi meetodit on võimalik kasutada ökonomeetrilise mudeli parameetrite hindamiseks, iga ettevõtja jaoks reservide kättenäitamiseks piima tootmiskulude vähendamiseks ning ettevõtjate majandusliku tegevuse hindamiseks (prognoosimiseks).