



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search  
<http://ageconsearch.umn.edu>  
[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

# **Using the information about dairy herd's genetic level and milk quality in explaining the technical efficiency of Estonian dairy farms: a two-stage (DEA and Tobit) approach**

Helis Luik<sup>1</sup>, Ants-Hannes Viira<sup>2</sup>, Rando Värnik<sup>2</sup>

<sup>1</sup> Estonian University of Life Sciences, Institute of Economics and Social Sciences, e-mail: helis.luik@emu.ee

<sup>2</sup> Estonian University of Life Sciences, Institute of Economics and Social Sciences



**Paper prepared for presentation at the EAAE 2014 Congress  
'Agri-Food and Rural Innovations for Healthier Societies'**

August 26 to 29, 2014  
Ljubljana, Slovenia

*Copyright 2014 by Helis Luik, Ants-Hannes Viira and Rando Värnik. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*

## Abstract

This paper investigates the effects of dairy herd's genetic level and milk quality on the technical efficiency (TE) of Estonian dairy farms in 2012. A two-stage approach was used that combined data envelopment analysis (DEA), for finding the TE scores, and Tobit regression, for estimating the effects of the chosen variables on farm TE. Our results indicate that relative breeding value for milk production, which is a measure of herd's genetic level, has a positive effect on farm TE. Milk quality affects TE also positively: higher somatic cell counts decreased and higher content of milk solids increased farm TE.

Keywords: technical efficiency, dairy farms, breeding values, milk quality, DEA.

## 1. Introduction

Dairy is one of the most important sectors in Estonian agriculture. In 2013, milk accounted for 27.2% of the value of agricultural output. During the past decade the productivity of dairy cows has increased substantially in Estonia. From 2003-2012 the average milk yield increased by 45.4% from 5176 to 7526 kg per cow (Statistics Estonia, 2014). This increase in yield could be associated with the investments in modern cowsheds, improved technologies and feeding (introduction of total mixed feeding rations). The accession to the EU in 2004 significantly increased the level of investment subsidies and direct payments compared to the previous years, which also helped in improving the access to credit (Estonian Ministry of Agriculture, 2012).

In Europe, average farm sizes have been increasing and total number of farms decreasing since 1950, the same trend is also visible in Estonia. From 2003 to 2012 the number of dairy farms under milk recording decreased by 69.3% from 2712 to 833. In the herd group of up to 100 dairy cows the reduction in herd numbers was 74.2% and in the group of herds with more than 100 cows the reduction of number of herds was 20.0% (Estonian Animal Recording Centre, 2004; 2013b). Thus, the sustainability of small farms a relevant issue in agricultural policy debates in Estonia.

The relationship between farm size and productivity is widely studied. Newman and Matthews (2006) found a positive relationship between the farm size and productivity in Irish dairy farms. The productivity growth was fastest in farms with greater than 100 ha and slowest on the farms with less than 30 ha. In addition, they found that TE was higher in larger farms. Bojnec and Latruffe (2013) found that farm size has a positive effect on TE in Slovenia in contrast to Rasmussen (2010) who reached the opposite results in Denmark: smaller farms have, on average, higher TE. Therefore, on the one hand, farm size affects productivity due to the scale effect, implying that the optimal size of farm is economically important. On the other hand, farm size is not so important – smaller (family) farms maintain the rural lifestyle and therefore, these small farms are socioeconomically important even though they may be technically inefficient. It is obvious that size affects efficiency but it is increasingly important to examine the factors that lead to inefficiency regardless of farm size.

Several other external and internal factors affect TE, such as weather, subsidies, health of the herd, farmer's age, education etc. The important indicators of assessing the dairy farms are connected with herd, such as production quantity and quality, breeding values, cost of the feed. Hansson (2007) found that milk yield per cow and the occurrence of mastitis are significant predictors of good economic performance. Hansson et al. (2011) assessed which preventive measures against mastitis can increase the TE because of the loss of revenue caused by mastitis, and revealed some crucial simple operations to be implemented. The genetics, more specifically various breeding values are other factors to consider in the TE analyses. Roibas and Alvarez

(2010; 2012), Steine et al. (2008) have found, based on the empirical analysis, that genetic progress can be considered a powerful tool for increasing farm profitability.

Thus, if one wants to forecast the productivity growth of dairy farms for the next decade in Estonia, it is reasonable to assume that the likelihood of attaining similar growth in milk yield (45.4%) in the next ten years is low due to relatively high average yield today. Therefore, dairy farmers, while looking for the sources of productivity and efficiency growth should pay attention to more detailed and nuanced factors that affect productivity of dairy farms.

Therefore, this paper aims to include the information about the genetic level of dairy herds and milk quality in the analysis of TE. In addition, more common factors like farm manager's age, farm size, farm labour and land use are considered. For that the 2012 Estonian farm level FADN data is combined with the 2012 herd level milk recording data. The Data Envelopment Analysis (DEA) is used for finding the TE scores for sample farms and Tobit regression is used to estimate the effects of various factors on the TE of dairy farms.

## 2. Data and method

The main focus of this study is on integrating the information about the genetic level of dairy herds and milk quality in the farm TE analysis. For this purpose two data sets were combined: Estonian farm accountancy (FADN) data (Rural Economy Research Centre, 2013) about farm accounts in 2012, and the data from Estonian Animal Recording Centre (2013a) regarding the herd level breeding values, quantity and quality of the milk produced, and herd replacement statistics.

The Estonian FADN (farm level) dataset was used to calculate the TE of Estonian dairy farms in 2012. FADN data provide information on production outputs, labour input, capital input, variable and fixed costs, agricultural area, number of cows and other relevant farm income and cost figures. In 2012, there were 657 farms in Estonian FADN sample of which 179 were specialized on milk production, 127 were specialized on grazing livestock, and 103 were mixed producers.

The farm level breeding values, milk production and quantity data, and herd replacement statistics were available for all Estonian dairy farms who participated in milk recording programme in 2012. In 2012 93.8% of Estonian dairy cows from 833 herds were under milk recording (Estonian Animal Recording Centre, 2013b)<sup>1</sup>. After integrating the FADN and milk recording datasets and removing the outliers, the data from 147 dairy producers remained valid for the analysis.

The most common methods of assessing the TE are DEA and Stochastic Frontier Analysis (SFA). Both of them have advantages and disadvantages. DEA is based on linear programming whereas SFA employs econometric techniques. DEA has several advantages: ability to accommodate a multi-input and multi-output models; the input and output values can be measured in different units; it takes into account returns to scale in calculating efficiency, allowing for the concept of increasing or decreasing efficiency based on size and output levels. One of the most important advantages of DEA is that it does not need any restriction on the functional form of the production relationship between inputs and outputs (Bogetoft and Otto, 2010). Considering the advantages of the DEA, this method is used in this paper.

Several TE analyses in the agricultural sector have used an approach, where in the first stage DEA is used to calculate the efficiency scores by using the physical and/or monetary quantities of inputs and outputs; and then in the second stage the estimates of the factors affecting the TE are found by using Tobit model (Davidova and Latruffe, 2007; Barnes, 2006). In this analysis a similar two-stage approach was used. First, the TE of dairy farms was

---

<sup>1</sup> Participation in milk recording is not compulsory.

estimated using the DEA. The second stage of the analysis aimed to estimate the effects of genetic information, milk quality and other farm-specific variables on the TE of dairy farms. For that purpose a Tobit regression was used.

## 2.1 Data Envelopment Analysis

Based on the productivity analysis it is possible to identify the most efficient farms. The higher the output per input, the higher production efficiency is. The production efficiency is easy to calculate if there is only one output and one input. However, usually there are no such production units which produce only one output with one input. The problem is that if there are several inputs and outputs which are measured in different ways, it is impossible to apply the simplest productivity measurement.

To overcome this obstacle, one option amongst the others is to apply nonparametric piecewise linear programming method DEA. The idea of the method was proposed by Farrell (1957). Accordingly, the TE reflects the ability of economic unit to obtain maximum output from a given set of inputs. Farrell defined the TE in two ways: the ability of a farm to produce the maximum feasible output with a given bundle of inputs or the ability of a farm to use minimum inputs to produce a given level of output (Cooper et al., 2004). DEA analysis can identify the efficient units, and results for inefficient units will show by how much each input can be reduced or output increased to produce an optimal output (Cooper et al., 2004).

The objective of performance evaluation is to evaluate the current farm operation internally and to benchmark it against similar farm operations externally in order to identify the best practice. Such best practice can be empirically identified and the efficient frontier can be empirically estimated based upon observations on one farm operation over time or similar farm operations at a specific time period (Zhu, 2009).

Each farm has a set of inputs and outputs, representing multiple performance measures. To evaluate the farm's TE input-oriented variable returns to scale model (VRS) was used in this analysis. To express mathematically the variable returns to scale model (1), we considered a set of  $n$  farms. Each farm uses  $m$  inputs to produce  $s$  outputs. Specifically, farm <sub>$j$</sub>  uses  $x_{ij}$  input  $i$  and produce  $y_{rj}$  output  $r$ . The TE measure under the assumption of variable return to scale can be formulated as:

$$\begin{aligned} \theta^* &= \min \theta & (1) \\ \text{subject to} \\ \sum_{j=1}^n x_{ij} \lambda_j &\leq \theta x_{i0} & i = 1, 2, \dots, m; \\ \sum_{j=1}^n y_{rj} \lambda_j &\geq y_{r0} & r = 1, 2, \dots, s; \\ \lambda_j &\geq 0 & j = 1, 2, \dots, n. \\ \sum_{j=1}^n \lambda_j &= 1 \end{aligned}$$

Where  $\theta$  is a scalar;  $x$  and  $y$  are input and output quantities respectively.  $\lambda$  is a vector describing the contribution of the benchmark farms to the virtual farm on the frontier. Using the variables  $\lambda$  and  $\theta$ , the model is solved once for each farm, looking for the largest radial

contraction of the input vector  $x_i$  within the technology set. The value of  $\theta^*$  obtained is the efficiency score for the  $n$ -th farm, with a value of 1 indicating a point on the frontier and hence a technically efficient farm (Zhu, 2009; Coelli, 2005).

Next the variables used in the DEA model are described. The input and output characteristics for the model are selected based on the dairy producers' specific inputs and outputs and are in line with the approach used by other researchers in the same field (e.g. Davidova and Latruffe, 2007; Rasmussen, 2010).

There are two outputs in the model and both of them are in monetary value. The first output is sales revenue of milk and dairy products ( $y_1$ ), the second aggregated measure contains sales revenue of live animals, other agricultural production and services provided for other farmers ( $y_2$ ). The model has six parameters as inputs: the number of dairy cows ( $x_1$ ) represents the annual average number of cows. Agricultural area ( $x_2$ ) is measured as the farm's total utilised agricultural area in hectares. Labour ( $x_3$ ) is in the form of annual working hours and includes all labour, both paid and unpaid, which has contributed to the work on the farm during the accounting year. Feed costs ( $x_4$ ) include both purchased and home grown feed. Other costs ( $x_5$ ) include: livestock and crop specific cost (excluding costs of the feed), energy and other direct inputs. The capital costs ( $x_6$ ) include depreciation of machinery and buildings, maintenance of machinery and building, insurance, interests, renting costs and contract work.

The descriptive statistics of the outputs and input used in the DEA<sup>2</sup> are presented in Table 1. Due to the dualistic farm structure in Estonia, the range of values of different variables is remarkable, for example the minimum number of dairy cows in the herd is 6 and the maximum is 1643. The average is 210 cows in the herd, but more important is that median is 87, which means that most of the herds are rather small. 68% (100 farms) of the sample herds had less than 200 cows, and just 3 farms had more than 1000 cows.

**Table 1. Statistics of the used outputs and inputs for DEA in 2012**

	Unit	Min	Max	Mean	St. Dev.	Median
<i>Outputs</i>						
Sales revenue of milk and dairy products ( $y_1$ )	Euro	4,204	4,051,552	470,019	693,867	157,897
Sales revenue of animals and other agricultural products ( $y_2$ )	Euro	1,664	6,377,602	226,522	581,546	53 404
<i>Inputs</i>						
Dairy cows ( $x_1$ )	Number	6	1,643	210	284	87
Agricultural area ( $x_2$ )	Ha	27	5,729	693	893	269
Labour ( $x_3$ )	Hours	2,000	254,376	29,611	38,759	10,800
Feed costs ( $x_4$ )	Euro	4,904	3,027,454	314,337	455,077	93,429
Other costs ( $x_5$ )	Euro	4,923	5,327,337	305,671	559,072	94,730
Costs of the capital ( $x_6$ )	Euro	4,538	2,217,614	208,233	307,683	78,155

## 2.2 Tobit regression

Before specifying the model for Tobit regression a one-way ANOVA test<sup>3</sup> on the differences of various characteristics for low, medium and high TE farms was performed. 147 sample farms were divided in groups of low, medium and high TE so that each group contained equal number (49) of farms. In low TE group the TE scores ranged from 0.549 to 0.836; in medium TE group the TE scores ranged from 0.840 to 0.968; and in the high TE group the TE scores ranged between 0.970 and 1.

In the second stage of TE analysis the following Tobit model<sup>4</sup> (2) was specified to estimate the effects of selected variables on the TE of Estonian dairy farms in 2012:

<sup>2</sup> Software package DEAP 2.1 was used for estimating the TE scores

<sup>3</sup> Software Statistica 12 was used for ANOVA

<sup>4</sup> R package censReg, version 0.5-20, was used for Tobit regression

$$TE\_VRS_i = \beta_0 + \beta_1 Cows\_50_i \times RBV\_milk_i + \beta_2 Cows\_50\_100_i \times RBV\_milk_i + \beta_3 Cows\_100\_300_i \times RBV\_milk_i + \beta_4 Cows\_300_i \times RBV\_milk_i + \beta_5 Age_i + \beta_6 EH\_share_i + \beta_7 EH\_share_i^2 + \beta_8 SCC_i + \beta_9 Yield_i \times Milk\_solids_i + \beta_{10} Own\_feed_i + \beta_{11} Own\_feed_i^2 + \beta_{12} Land\_cow_i \times Grassland\_share_i + \varepsilon_i \quad (2)$$

Where  $TE\_VRS_i$  is the variable returns to scale TE score of farm  $i$ ,  $Cows\_50_i$ ,  $Cows\_50\_100_i$ ,  $Cows\_100\_300_i$  and  $Cows\_300_i$  are dummy variables indicating the size of the dairy farm measured in number of cows: <50 cows, 50-<100 cows and 100-<300 cows and  $\geq 300$  cows. Due to the positive correlation between farm size and relative breeding values for milk ( $RBV\_milk_i$ ), interactions between farm size group dummy variables and  $RBV\_milk_i$  were used in the model (2).  $RBV\_milk_i$  indicates the difference between the herd's average relative breeding value for milk production and the respective national average. Thus, the  $RBV\_milk_i$  indicates by how many units the herd's respective breeding value deviates from the national average. Due to the fact that the breeding values for the breeds Estonian Holstein (EH) and Estonian Red (ER) are calculated separately, the farm level figure is weighed using the share of EH and ER dairy cows in the herd.  $Age_i$  indicates the age of the farm manager.  $EH\_share_i$  indicates the proportion of EH cows in the herd.  $SCC_i$  indicates the average level of somatic cell count at the herd level in 2012. This is a proxy of milk quality (hygiene).  $Yield_i$  indicates the herd average milk yield per dairy cow.  $Milk\_solids_i$  indicates the aggregated milk fat and protein content. This is another factor describing the milk quality. Due to the negative correlation between milk yield and milk solids, an interaction of  $Yield_i$  and  $Milk\_solids_i$  was used in the model (2).  $Own\_feed_i$  indicates the proportion of farm-grown feed in the total feed costs.  $Land\_cow_i$  indicates the hectares of agricultural land per dairy cow and is an indicator of land use intensity in the farm.  $Cereal\_share_i$  measures the proportion of land allocated for cereal and oilseeds production.

### 3. Results and discussion

#### 3.1 Technical efficiency scores

The first step in the analysis involved computing the TE for each farm. The TE scores were calculated by using the input-oriented VRS model of given output and input matrix. The VRS value equal to 1.00 (100%) means that the farm is technically efficient and it is producing the maximum outputs per given inputs. The results indicate that 29% of the farms are technically efficient. 39% of the total sample has VRS score more than 0.95, which denotes that input usage should be reduced less than 5% in these farms in average. The most critical group contains 10% of the total sample: those producers should decrease inputs more than 30%.

**Table 2. Distribution of technical efficiency scores in variable returns to scale model**

VRS technical efficiency score	Number of farms	Percentage of farms	Cumulative percentage
<0.70	14	10	10
0.70-0.79	22	15	24
0.80-0.84	14	10	34
0.85-0.89	16	11	45
0.90-0.94	24	16	61
0.95-0.99	14	10	71
1.00	43	29	100

The scale efficiency results reflect that 18% of dairy farms were operating at the optimal scale. 18% of farms were operating at decreasing return to scale and 64% of farms were

operating at increasing return scale. Based on this, it can be concluded that there are more than half of dairy producers who can increase their TE by increasing the size of their production.

### 3.2 Analysis of variance (ANOVA)

In Table 3 the results of ANOVA tests are reported. It occurs that the average number of dairy cows per heard does not differ significantly in low, medium and high TE groups. While in average farm operators were younger in higher TE groups, the differences were not statistically significant. In medium and high TE groups the average share of EH cows in herd's total was significantly higher, compared to the low TE group.

The main purpose of this paper was to include data on herd's genetic level and milk quality in the TE analysis. According to the results in Table 3, the relative breeding value for milk is significantly positively ( $p < 0.01$ ) related to farm's TE. In high TE farms, the average value of *RBV\_milk* was -0.686, while in medium TE farms it was -2.442 and on low TE group it was -3.593. While the positive relationship between genetic level and farm TE may be related to higher productivity dairy cows, one could argue that breeding values are a result of conscious selection of bulls which sperm is used for insemination. One could assume that if the bulls are selected consciously, the same level of consciousness is used in other aspects of everyday farm management. Therefore, the high *RBV\_milk* could also be regarded as an indicator of farm management.

**Table 3. Results of one-way ANOVA on TE scores**

Variable	Technical efficiency			F-value	P-value
	Low	Medium	High		
Number of farms	49	49	49	-	-
<i>TE_vrs</i>	0.736	0.905	0.997	401.40	0.000***
<i>Cows</i>	178.0	212.8	238.5	0.555	0.575
<i>Age</i>	54.0	51.2	49.4	1.947	0.146
<i>EH_share</i>	0.646	0.808	0.811	3.942	0.022**
<i>RBV_milk</i>	-3.593	-2.442	-0.686	5.001	0.008***
<i>SCC</i>	406.0	317.4	321.1	5.974	0.003***
<i>Fat</i>	4.219	4.081	4.073	5.390	0.006***
<i>Protein</i>	3.329	3.360	3.357	1.600	0.205
<i>Yield</i>	6,413	7,794	7,764	18.406	0.000***
<i>Own_feed</i>	0.733	0.671	0.613	5.371	0.006***
<i>Land_cow</i>	4.439	3.663	3.566	3.874	0.023**
<i>Grassland_share</i>	0.715	0.663	0.698	0.620	0.539
<i>Sales_milk</i>	1,637	2,058	2,105	16.017	0.000***
<i>Sales_other</i>	628	938	1,004	3.043	0.051*
<i>Labour_hours</i>	166.6	155.2	162.1	0.302	0.740
<i>Feed_costs</i>	1,302	1,325	1,320	0.047	0.954
<i>Other_costs</i>	1,167	1,311	1,275	0.761	0.469
<i>Capital_costs</i>	869	1,049	992	3,856	0.023**
<i>Milk_sold_share</i>	0.916	0.929	0.944	3.430	0.035**
<i>Milk_price</i>	0.286	0.292	0.288	0.720	0.488
<i>Cow_age</i>	79.3	71.5	73.7	2.881	0.059*
<i>Milk_day</i>	9.77	11.45	11.83	14.80	0.000***
<i>Culling</i>	0.262	0.275	0.271	0.177	0.838

\*Significant at 0.1 level; \*\*Significant at 0.05 level; \*\*\*Significant at 0.01 level

Two indicators of milk quality were used in the analysis. Somatic cell count could be regarded as an indicator of hygiene. From Table 3 it stems that in low TE farms the *SCC* was significantly ( $p < 0.01$ ) higher compared to medium and high TE farm groups. Higher *SCC* indicate poor hygiene and day-to-day farm management. Therefore, our results imply that *SCC* could also be regarded as an indicator of farm performance. At the same time higher *SCC* values may imply problems with milk quality, which could mean that lower percentage of milk from

the total milk production might be sold, and the milk price could be lower compared to farms with low SCC. From Table 3 it can be seen that in low TE farms the percentage of milk sold out of total milk production is lower than in medium and high TE farm groups ( $p < 0.05$ ), however, there are no significant differences in average milk price.

It appears that in medium and high TE farms, the average fat content of milk is lower ( $p < 0.01$ ). This is related to the average milk yield – with increasing yield the fat content decreases. However, the average protein content of milk did not differ significantly in various farm TE groups. The effects of milk solids on farm TE is further analysed in the Tobit regression.

From Table 3 it appears that another indicator that is strongly related to farm TE is average milk yield per cow's lifetime (kg/day). In high TE farms this figure was higher compared to medium and low TE farms. This figure contains information about cow's average yield, age of the first calving, calving interval, lactation curve, age when culled etc. Therefore, milk per day of life could be regarded as a performance indicator of dairy farms that could be obtained from the milk recording data.

### 3.3 Tobit regression

The estimated coefficients and average marginal effects of the explanatory variables in the specified Tobit model (2) are presented in Table 4. The parameter estimates indicate that the genetic level of dairy herd, indicated by *RBV\_milk* has a positive effect on dairy farms TE in farm size groups 50-<100, 100-<300 and  $\geq 300$  dairy cows. In size group of <50 dairy cows, the effect of *RBV\_milk* on farm TE was negative. The results imply that the genetic level has stronger positive effect on farm TE in larger herds. From 1997-2010 the Estonian average *RBV\_milk* for EH cows increased by 27 units, i.e. by 2.1 units per year. This implies that the TE of dairy farms could be increased via breeding, but the effects are not immediate.

Age of the farm manager had a significant ( $p < 0.1$ ) negative effect on the TE in dairy farms. If farms with managers of 50 and 60 years of age are compared, the average TE in farms with older managers would be lower by 0.012 points. Hence, while the effect is statistically significant, its magnitude is relatively low.

The relationship between the shares of EH cows in the dairy herds and the TE of dairy farms was found nonlinear. First it increases with increasing share of EH cows in the herd, and starts to decrease after the share of EH cows in the herd increases 65%.

The results in Table 4 indicate that milk quality has positive effects on farm TE. Milk quality is reflected by two variables: somatic cell count (SCC) and interaction of milk yield per cow and milk solids. The results indicate that increasing number of average SCC has negative effect on TE. If SCC increases by 100, the average TE decreases by 0.02 points. In case of milk solids, the interaction with milk yield was used due to the correlation between these factors. Increasing yields imply decreasing average milk solids. The results from Table 4 indicate that if milk yield is constant, then increasing milk solids have significant ( $p < 0.05$ ) positive effect on the TE of dairy farms and the effect is stronger in case of higher yields.

It was also analysed to what extent the production of all the feed at the farm affects the TE. The results indicate that higher proportion of farm-grown feed in total feed costs has a negative effect on TE. However, the relationship between the share of *Own\_feed* and TE is nonlinear, and if the proportion of the farm-grown feed in total feed costs exceeds 76%, the negative effect is reduced. However, this implies that in Estonia the feeding strategy that involves usage of bought concentrates is positively affecting the TE.

In Estonia the land use in dairy farms is relatively extensive, especially if it is compared to the Central European countries. The results in Table 4 indicate that extensive land use (relative to the number of dairy cows) could have negative effect on the farm TE if the proportion of

grassland on total land use is increased. Our results show that if there are e.g. 3.6 ha of agricultural land per cow, then increasing the share of land allocated for the grassland from 60% to 70% increases the average TE by 0.005 units.

**Table 4. The results of the Tobit estimates**

Variables	Estimate	Std. error	Marginal effect	Std. error
<i>Intercept</i>	1.3060***	0.0188	-	-
<i>Cows_50 x RBV_milk</i>	-0.0077**	0.0033	-0.0062**	0.0026
<i>Cows_50_100 x RBV_milk</i>	0.0095*	0.0057	0.0076	0.0046
<i>Cows_100_300 x RBV_milk</i>	0.0230***	0.0076	0.0184***	0.0060
<i>Cows_300 x RBV_milk</i>	0.0109**	0.0055	0.0087*	0.0044
<i>Age</i>	-0.0015*	0.0009	-0.0012	0.0007
<i>EH_share</i>	0.2921**	0.1396	0.2336**	0.1116
<i>EH_share</i> <sup>2</sup>	-0.2269*	0.1229	-0.1814*	0.0982
<i>SCC</i>	-0.0002***	0.0001	-0.0002**	0.0001
<i>Yield x Milk_solids</i>	0.0003**	0.0001	0.0003**	0.0001
<i>Own_feed</i>	-1.2980***	0.4404	-1.0381***	0.3473
<i>Own_feed</i> <sup>2</sup>	0.8508***	0.3259	0.6804***	0.2578
<i>Land_cow x Grassland_share</i>	-2.1920**	0.0082	-0.0130*	0.0066
<i>Log-Likelihood</i>	45.850			
<i>Number of observations</i>	147			

\*Significant at 0.1 level; \*\*Significant at 0.05 level; \*\*\*Significant at 0.01 level

#### 4. Conclusions

Farm's technical efficiency is an important attribute of farm's competitiveness compared to other farms in the sector. In this paper the TE of Estonian dairy producers was analysed. In addition to farm-specific socioeconomic variables the effects of a herd's genetic level (relative breeding value for milk production) and milk quality attributes (somatic cell count, milk solids) on the TE of Estonian dairy farms in 2012 were analysed.

For the analysis, farm level accounts data from Estonian FADN and herd level genetic information and milk quality data were merged. TE scores were found with DEA method. ANOVA was used to test the differences of various characteristics for low, medium and high TE farms. Comparing the means of the characteristics low, medium and high TE farms helped in detecting interrelations between TE scores and different explanatory variables. In the second stage of TE analysis a Tobit regression was used to estimate the effects of explanatory variables on the farm level TE.

Our results indicate that herd's genetic level affects positively the TE of dairy farm. This is in line with the findings of Roibas and Alvarez (2010; 2012), and Steine et al. (2008). Also, it appeared that milk quality is significantly affecting the TE. Lower levels of somatic cell count and higher levels of milk solids had a positive effect on farm TE.

Usage of farm-grown feed had a non-linear effect on farm TE. It appeared that if the share of own-grown feed increases, the TE is decreasing. However, if the share of own grown feed exceeds 76%, the TE starts to increase. Also, it appeared that extensive land could decrease the TE of dairy farms if this land is used for grasslands. If the share of grasslands on total land use is high, the extensive land use decreases the farm's TE compared to other farms.

This analysis showed that integrating farm accounts data, herd level genetic information, and milk quality attributes enables to use more specific factors to explain the variation of TE between dairy farms. Considering the already relatively high milk yields in Estonian dairy herds, more and more nuanced factors should be considered in analysing and improving the productivity and efficiency of dairy farms.

## 5. Acknowledgements

The authors would like to thank Estonian Rural Economy Research Centre for providing the farm level FADN data, and Estonian Animal Recording Centre for providing the herd level data about the genetics, milk quality and herd replacement statistics, professor Haldja Viinalass and associate professor Tanel Kaart for the consultations on breeding values.

## 6. References

- Barnes, A.P. (2006). Does Multi-Functionality Affect Technical Efficiency? A Non-Parametric Analysis of the Scottish Dairy Industry. *Journal of Environmental Management* 80(4): 287-294.
- Bogetoft, P., Otto, L. (2010) *Benchmarking with DEA, SFA, and R*. New York: Springer, 351 p.
- Bojnec, Š., Latruffe, L. (2013) Farm size, agricultural subsidies and farm performance in Slovenia. *Land Use Policy* 32: 207–217.
- Cooper, W.W., Seiford, L.M. and Zhu, J. (2004) *Handbook on Data Envelopment Analysis*. Springer. 608 p.
- Coelli, T. J., Rao D.S., P. O' Donnell C. J., Battese G. E. (2005) *An Introduction to Efficiency and Productivity Analysis*. Springer, 350 p.
- Davidova S., Latruffe L. (2007) Relationship between technical efficiency and farm financial management: the example of the Czech Republic *Journal of Agricultural Economics*, 58, pp. 269–288
- Estonian Animal Recording Centre (2013a) Data provided to Estonian University of Life Sciences.
- Estonian Animal Recording Centre (2013b) *Results of Animal Recording in Estonia 2012*. Elmatar, 2013, 52 p.
- Estonian Animal Recording Centre (2004) *Results of Animal Recording in Estonia 2003*. 45 p, [https://www.jkkeskus.ee/assets/tekstid/aastaraamatud/aastaraamat\\_2003.pdf](https://www.jkkeskus.ee/assets/tekstid/aastaraamatud/aastaraamat_2003.pdf) (last accessed 30.06.2014)
- Estonian Ministry of Agriculture (2012) *Estonian dairy strategy 2012-2020*. (in Estonian), <http://www.agri.ee/public/juurkataloog/ARENDUSTEGEVUS/piimandusstrateegia-2012-2020.pdf> (last accessed 31.01.2014)
- Farrell, M.J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*. Vol. 120 (3): 253-290.
- Hansson, H. (2007). The links between management's critical success factors and farm level economic performance on dairy farms in Sweden. *Food Economics-Acta Agricult Scand C*. 4: 77-88
- Hansson, H. Szczesna-Rundberg, M., Nielsen, C. (2011) Which preventive measures against mastitis can increase technical efficiency of dairy farms? *Animal: An International Journal of Animal Bioscience*. 5: 632-640.
- Newman, C., Matthews, A. (2006) The productivity performance of Irish dairy farms 1984–2000: a multiple output distance function approach. *Journal of Productivity Analysis*, Springer, vol. 26(2), pp 191-205.
- Rasmussen, S. (2010). Scale efficiency in Danish agriculture: an input distance-function approach. *European Review of Agricultural Economics*, Foundation for the European Review of Agricultural Economics, vol. 37(3), pp 335-367.
- Roibas, D., Alvarez, A. (2010) The impact of genetic progress on the profits of dairy farmers. *Journal of Dairy Science* (93) 4366–4373.

- Roibas, D., Alvarez, A. (2012) The contribution of genetics to milk composition: evidence from Spain. *Agricultural Economics* (43) 133-141.
- Rural Economy Research Centre (2013) Farm level FADN data provided to Estonian University of Life Sciences.
- Steine, G., Kristoffersson, D., Guttormsen, A. (2008) Economic Evaluation of the Breeding Goal for Norwegian Red Dairy Cattle. *Journal of Dairy Science* (91) 418-426.
- Statistics Estonia (2014) Online statistical database at [www.stat.ee](http://www.stat.ee) (last accessed 31.01.2014)
- Zhu, J. 2009. *Quantitative Model for Performance Evaluating and Benchmarking*. USA: Springer, 327 p.

## Annex I Definitions and descriptive statistics of variables used in ANOVA and Tobit regression

<b>Variable</b>	<b>Definition and measurement</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>St. Dev.</b>	<b>Median</b>
<i>TE_vrs</i>	TE score estimated in the 1st stage of the analysis	0.549	1.000	0.879	0.118	0.918
<i>Cows</i>	Annual average number of dairy cows	6,0	1,643	210	284	87
<i>Age</i>	Age of the farm manager in 2012, years	20	79	51.6	11.7	52.0
<i>EH_share</i>	Share of Estonian Holstein cows in total number of cows	0.000	1.000	0.754	0.334	0.938
<i>RBV_milk</i>	Difference between weighted (according to the share of breeds) farm's average relative breeding value for milk and national average breeding value for milk production.	-18.700	7.609	-2.241	4.706	-1.700
<i>SCC</i>	Farm average somatic cell count, SCC thousand/ml	97	918	348	148	313
<i>Fat</i>	Farm average fat content in milk	3.320	5.330	4.124	0.255	4.120
<i>Protein</i>	Farm average protein content in milk	3.080	3.620	3.349	0.097	3.340
<i>Yield</i>	Milk yield of dairy cows, kg/cow	4,225	9,953	7,323	1,432	7,275
<i>Own_feed</i>	Proportion of farm-grown feed in total feed costs	0.184	1.000	0.672	0.187	0.696
<i>Grassland_share</i>	Proportion of grazing area in total agricultural area	0.222	1.000	0.692	0.233	0.708
<i>Sales_milk</i>	Sales revenue from milk per dairy cow, euros/cow.	701	2,891	1,933	495	1,947
<i>Sales_other</i>	Sales revenue from other activities per dairy cow, euros/cow	82.9	4,070	857	818	573
<i>Land_cow</i>	Agricultural land per dairy cow, ha/cow	1.093	9.946	3.889	1.735	3.488
<i>Labour_hours</i>	Labour hours per dairy cow, hours/cow	51.5	397	161	72.8	137
<i>Feed_costs</i>	Feed costs per dairy cow, euros/cow	418	2,323	1,316	392	1,308
<i>Other_costs</i>	Other costs per dairy cow, euros/cow	366	3,263	1,251	604	1,134
<i>Capital_costs</i>	Capital costs per dairy cow, euros/cow	302	1,920	970	334	977
<i>Milk_sold_share</i>	Amount of sold milk divided by total amount of milk produced	0.643	1.000	0.930	0.055	0.945
<i>Milk_price</i>	Value of milk sold, euros/kg	0.181	0.368	0.289	0.025	0.295
<i>Cow_age</i>	Average age of dairy cow when culled, months	36.5	115.9	72.8	15.4	70.4
<i>Milk_day</i>	Total milk production of a dairy cow divided by its lifetime in days, kg/day	4.89	15.65	11.02	2.17	10.96
<i>Culling</i>	Number of culled cows divided by annual average number of cows	0.038	0.595	0.270	0.108	0.263
N = 147						