**WAGENINGEN UNIVERSITY AND RESEARCH CENTER**

**DEPARTMENT OF SOCIAL SCIENCES**

**BUSINESS ECONOMICS GROUP**

# M.SC. MAJOR THESIS

# Comparing dairy farm efficiency using DEA and SFA

| | | |
|---|---|---|
| **Student:** | Krisztián Kovács | 811012-471-100 |
| **Supervisors:** | Dr. Grigorios Emvalomatis | Wageningen University (The Netherlands) |
| | Ir. Gerard Giesen | Wageningen University (The Netherlands) |
| | Dr. István Szűcs | University of Debrecen (Hungary) |
| **Study load:** | 36 Credits | |
| **Period:** | January 2010 - August 2010. | |
| **Code number:** | BEC-80436 | |

**WAGENINGEN**

2010

# Preface

Spring 2008 was the time when my Hungarian supervisor István Szűcs mentioned a university in the Netherlands where I can improve my knowledge about the Business Economics. So we decide that I will go to Wageningen as a double degree student and learn something more something special and interesting at Wageningen University. The nice international atmosphere and several sport opportunities made me attached to Wageningen for a long time.

As a Research Master Variant student after the minor thesis I had to start my major thesis really early of my study, so I made an appointment with Prof Dr Alfons Oude Lansink the chair of the Business Economics Group (BEC), who welcomed me as a colleague and introduced Grigorios Emvalomatis for me. That time he was one of the best supervisors for me to start my research about efficiency.

Grigorios I would like to thank you for your kind hospitality since we met, thank you for your patience about my schedule and my mistakes. I am very grateful for that. I know that it takes a lot of time for you to review my papers, thanks you for it. I learned a lot from your useful remarks. Now I have clearer view how I can measure efficiency with multiple inputs and outputs, thank to you. I cannot express how much I am grateful that you helped me to do my thesis. Thank you a lot (Σας ευχαριστώ πολύ!).

I would like to thank also Gerard Giesen from the BEC, who gave lots of advice how to implement and improve my thesis. (Dank u zeer!).

I would like to thank to István Szűcs, who is my friend and my Hungarian supervisor, who gave me a chance to continue my studies at international level and gave some useful advices about my work (Nagyon köszönöm!).

I would like to express my gratitude to all of my friends here in Wageningen and in Debrecen, especially, Gabriella Ujhelyi, Zoltán Nagy, Annamária Németh, Andrea Kovács, Dániel Oláh, Garbriella Vadinszki, Beáta Bittner, Erika Szabó and for my girlfriend Monika F. Kruger, who were always there for me and gave me some extra positive energy during these times. (Nagyon köszönöm! Muito Obrigado!)

Finally I would like to give an enormous thank for my family, especially my parents, Jolán Bózsvári and Béla Kovács and my sister Szilvia Kovács who have always been there for me and who supported me. Kedves Édesanyám, Édesapám és Hugi! Köszönöm, hogy mindvégig mellettem voltatok életem jó és a nehéz óráiban. Nagyon köszönöm a tőletek kapott szeretetet, bíztatást és támogatást, ami mindig átlendített a nehéz időszakokon. Nagyon szeretlek titeket!

There are so many people I could not list here, but I would like to say thank you for your contribution about my work and about my life.

THANK YOU VERY MUCH!

# Abstract

The abolishment of the dairy quota system in the EU is expected to increase competition across dairy farms in Europe. Assuming a common price for milk in the EU, only the most efficient farms will survive in the new environment. The main objective of the research is to compare dairy farms in Germany, The Netherlands and Hungary about their technical efficiency. In the first part of the research, the efficiency is measured by partial efficiency indexes using one dimensional efficiency measuring. In the second part, the Data Envelopment Analysis (DEA) and the Stochastic Frontier Analysis (SFA) have to be used to measure efficiency in a multidimensional space, using six inputs and two outputs. Both methods used for measuring the same technical efficiency score, but under different assumptions, thus the second objective was to compare not only countries but, methods results as well.

It appears from the results that the highest efficiency farms are in the Netherlands, and then Germany and Hungary follow. But the country results from the two methods are sometimes showing some difference, what might cause the different method assumptions and the different sample size as well. In our research the difference between the two methods was 10-15 percent for Germany and The Netherlands and because of the small sample size 30-35 percent for Hungary. If we want to eliminate the low sample size effect, we can assume a common frontier, which decreases the efficiency scores a bit, and makes the Hungarian results more reliable.

If we combine the two methods and the two kinds of frontiers results and take the averages of the two methods and the two kinds of frontier, we get that the most efficient farms are in the Netherlands with 87-88% efficient. The German farms are 72-80% efficient. The Hungarian farms are 73-79% efficient.

With respect the abolishment of the dairy quota system, our results suggest that the Dutch farms are the most efficient, thus probably they will increase their production after the quota system. But because the size of the country we cannot expect dramatic changes in the European Dairy market. The Germans farms efficiency is lower, but their efficiency is also lower, so we won't expect high increase about the dairy supply. The Hungarian dairy sector is not so efficient like the Dutch, and the size of the sector has also small among the European countries, thus if they want to survive the quota system demolishing, they have to increase their technical efficiency.

# Table of contents

# 1 Introduction

## 1.1 Background and research problem

The world milk production shows a continuous rising trend since 1961. In 2005 the world total fresh milk production was 541 million tonnes (FAOSTAT 2010). Since the introduction of milk quotas in 1984 the European Union (EU) production has stagnated around 149 million tonnes (EUROSTAT 2010). The milk quota system was introduced to stop over-production in Europe.

The biggest milk producer in the world is Europe (37.08%) including the European Union (26.22%). The second largest milk producer is the American continent (North-, Central-, South America and the Caribbean) which represents 28.65% of the total milk production in the world (FAOSTAT 2010). The biggest milk producer in the EU is Germany (18.98%), the second is France (16.13%), and the third is the United Kingdom (9.83%). The Netherlands and Hungary account for 7.31% and 1.22% of total EU production, respectively (FAOSTAT 2010). Currently, dairy farms in a given EU country are expected to be more or less competitive when compared to dairy farms in other countries. A reason for that is the quota system, which does not allow trading between countries, may protect farmers from international competition. Given that the quota system will be abolished in 2013, this will put pressure on less competitive farms in different countries. The issue of optimal use of resources becomes important.

As noted by Bauer et al. (1998), policy makers are particularly interested in the potential impact of their decisions on performance of firms. A firm that is inefficient is wasting inputs because it does not produce the maximum attainable output, given the quantity of inputs used, and hence the possibility of reducing average costs. Irrespective of whether a developed or developing economy is under consideration, findings from the study of technical efficiency have far-reaching policy implications.

Studying farm efficiency and the potential sources of inefficiency are therefore important from a practical and a policy point of view. On the one hand, farmers could use this information to improve their performance. On the other hand, policymakers could use this knowledge to identify and target public interventions to improve farm productivity and farm income (Solís et al., 2009).

This research focuses on estimating and comparing the levels of technical efficiency (TE) among Dutch, German and Hungarian dairy farms. The estimation of technical efficiency is carried using Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). The results produced by the two methods will be compared, because the two methods have different assumptions. However, they measure the same thing: farm efficiency in our case. The DEA and the SFA look at the problem from two perspectives, which validate the final results.

## 1.2 Objective and research questions

The first objective of the research is to measure dairy farms efficiency in Hungary, Germany and The Netherlands. Based on the results, we can assess the potential of dairy farms in the three countries to survive of the abolishment of the dairy quota system.The second objective is to compare parametric and non-parametric methods of efficiency measurement in practise. The research questions of this thesis are: What are the differences and the similarities in the Dutch, German and Hungarian dairy sectors? The dairy farms in which country (the Netherlands, Hungary or Germany) are more efficient compared to their national frontier? Do the results obtained by the two methods (DEA and SFA) differ substantially and if so, what could be the reason?

## 1.3 Outline of the research

A literature study is performed in two directions. Firstly, literature on the overviews of the world, EU, Dutch, German and Hungarian dairy farming is studied. Secondly, the efficiency measurement techniques in the dairy sector are reviewed.

The next step is the determination of the three countries dairy farm criteria to define what a dairy farm is. Because in most cases dairy farms produce more than one product, we need to define a rule to decide what constitutes a dairy farm. In other words, we need to decide what type of farms will be studied, i.e., specialised, diversified etc. For the analysis we select those farms which has 75% of the revenues coming from the milk producing activity and build up our panel database from 2001 to 2005. These data are available at different sources but mainly the FADN database. For the country overview following database are used: FAOSTAT, EUROSTAT.

To study the determinants of technical efficiency we use **data envelopment analysis (DEA)**, which is a non-parametric approach to the estimation of frontier and the calculation of efficiency measures (e.g., Tauer, 1998; Jaforullah and Whiteman, 1999; Stokes et al., 2007). The other approach that we use and compare with the previous approach is the **stochastic** (production, cost, or profit**) frontier analysis (SFA)** (e.g., Heshmati and Kumbhakar, 1994; Bravo-Ureta et al., 2008) which is an alternative parametric approach for the estimation of frontier functions using econometric techniques. It has advantages over DEA when data noise is a problem. These two methodologies have also been used to analyze the potential sources of inefficiency (e.g., Lawson et al., 2004; Tauer and Mishra, 2006; Murova and Chidmi, 2009). However, Kumbhakar and Lovell (2000) argue that a stochastic frontier model seems to be the most appropriate approach in studies related to the agricultural sector because of its ability to deal with stochastic noise, accommodate traditional hypothesis testing, and allow for single-step estimation of the inefficiency effects (Cabrera, 2010).

The rest of this thesis is organised as follows: Chapter 2 includes a literature review about main indicators of the World, European Union, Dutch, German and Hungarian dairy sector. These sections focused mainly on milk production and partial efficiency indexes. The second parts of this chapter introduce the technical efficiency measuring models and compare them. Chapter 3 contain the database, the empirical specification of the Data Envelopment Analysis and the Stochastic Frontier Analysis are introduced. Chapter 4 presents the result from the two models and compare them. The last chapter discuss the results and conclude them and give advices for the future research.

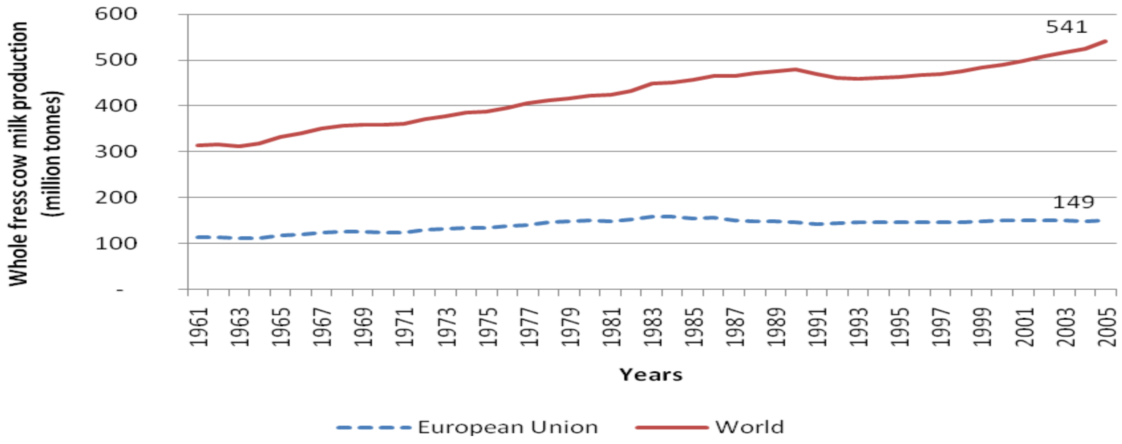
# 2 Literature review

This chapter presents a literature review about the main tendencies about dairy farming. First, the current trends in the World and the European dairy sector are introduced. These sections focus mainly on milk production, livestock and price trends. Next, differences between the German, Dutch and Hungarian farming conditions are explained. This refers to: livestock, production and price of milk. The final part of this chapter introduces and compares two techniques which can be used to measure dairy farm's efficiency.

## 2.1 Indicators of the World and the European Union dairy sector

The entire cattle population in the world in 2005 was approximately 1 372 million heads (FAOSTAT 2010). The biggest cattle livestock raising region was the American continent (503 million cattle; 37%), but most of the cattle was for beef production and not for dairy cows.

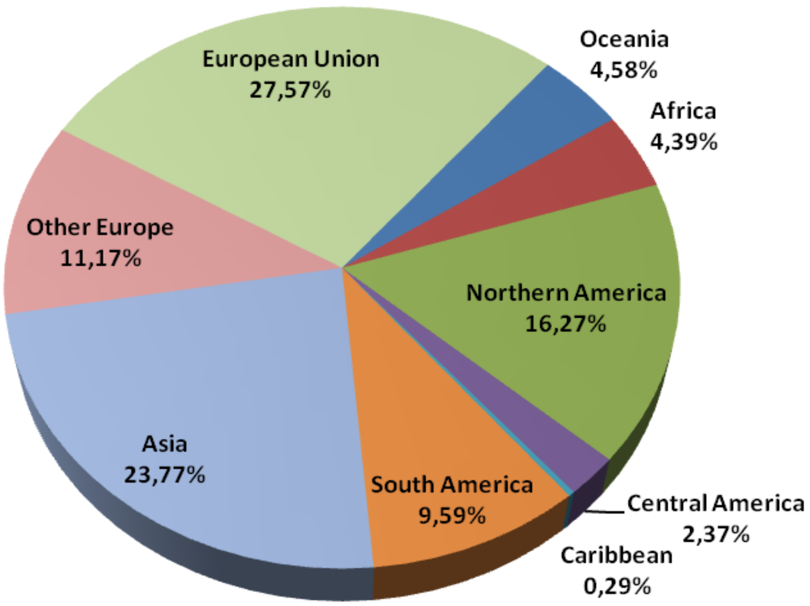
The world milk production shows a continuous rising trend since 1961. In 2005 the world total fresh milk production was 541 million tonnes (Figure 2.1). Since the introduction of milk quotas in 1984 the EU production stagnated at around 149 million tonnes. The milk quota system was introduced to stop over-production in Europe. Since 1984 there have been further reductions in quota of around 9%. The world milk production in 2005 is 541.34 million tonnes (Figure 2.1), of which the EU 25 was 149.26 million tonnes (FAOSTAT 2010).




Source: FAOSTAT 2010

*Figure 2.1: The World and the European Union cow milk production (whole, fresh) from 1961 to 2005*

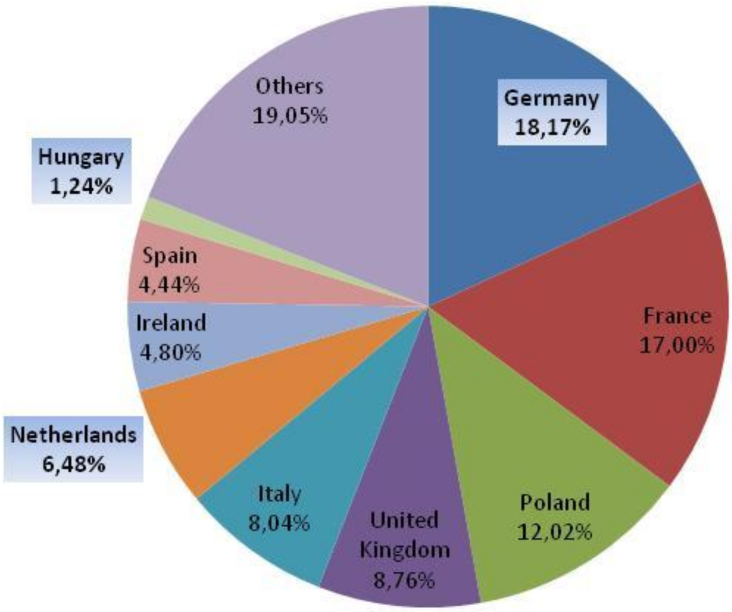
If we look at Figure 2.2, we can observe that the biggest milk producer in the world is the European continent (38.74%), including the European Union (27.57%), the second largest milk producer is the American continent (North-, Central-, South America and the Caribbean) which represents 28.52% of the total milk production in the world. The differences could be attributed to the size of the continent, but this may not be the only reason. The most prominent factors are: how many resources are available for milk production and how efficiently are these resources used. Another factor that is really important is the government policies in the different continents. We already mentioned the European milk quota, which restricts production or the subsidies connected to milk production.




Source: FAOSTAT 2010

*Figure 2.2: The Share of the World cow milk production in different continents in 2005.*

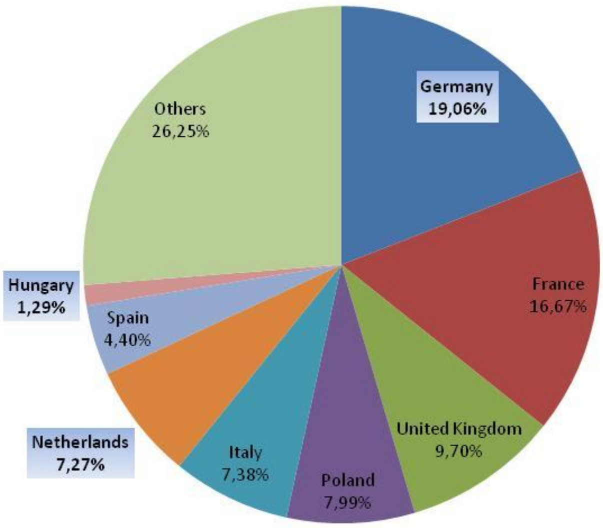
The total EU dairy cow population is 22.92 million heads. The biggest dairy livestock placed in Germany, 4.16 million dairy cows (Figure 2.3), which presents 21.15% of the whole European dairy livestock. Other big dairy raising countries are France (17.00%) and Poland (12.02%). The Netherlands and Hungary present 6.48% and 1.24% respectively.




Source: EUROSTAT 2010

*Figure 2.3: The Share of the European Union dairy stock in the 2005.*

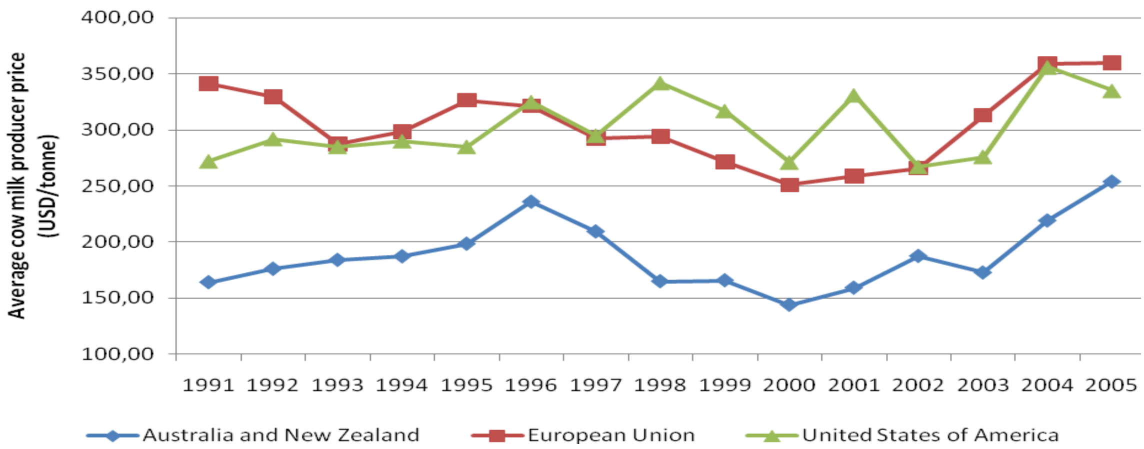
As we can see in the previous Figures (Figure 2.3) the biggest dairy livestock keeping countries are Germany, France and the Poland, but if we see the ratio of these countrie's livestock and milk production, we can be surprised. The biggest milk producer in the EU is Germany (19.06%), the second (Figure 2.4) is France (16.67%), but in the third place is the United Kingdom, and not Poland. The reason for this is that some countries use livestock-intensive technologies rather the livestock extensive technologies, which refer perhaps better production efficiency. For example the generally accepted productivity index is the average milk production per cow in the UK is 7261 kg/dairy cow, contrast with the polish 4336 kg/dairy cow. The Netherlands and Hungary milk production presents 7.27% and 1.29% respectively of the whole European fresh milk production.




Source: EUROSTAT 2010

*Figure 2.4: The European Union dairy milk production share in 2005.*

Efficient use of resources and high productivity can help to passes the potential of dairy farm to survive by reducing production cost. On the other hand, profitability is also affected by the price of the output. Therefore the second main focus of this section is the price tendencies and price fluctuation in the world and in the European market. Figure 1.5, represents average cow milk prices from 1991 to 2005 in the European Union, which is a restricted market, the United States, where some milk support program are in place and the Australian and New Zealand unregulated markets. Until 1991 the World market price (which is equal to the Australian and New Zealand milk price) is on average 30-40% lower than the restricted European milk price. The USA milk price more or less follows the European tendencies.

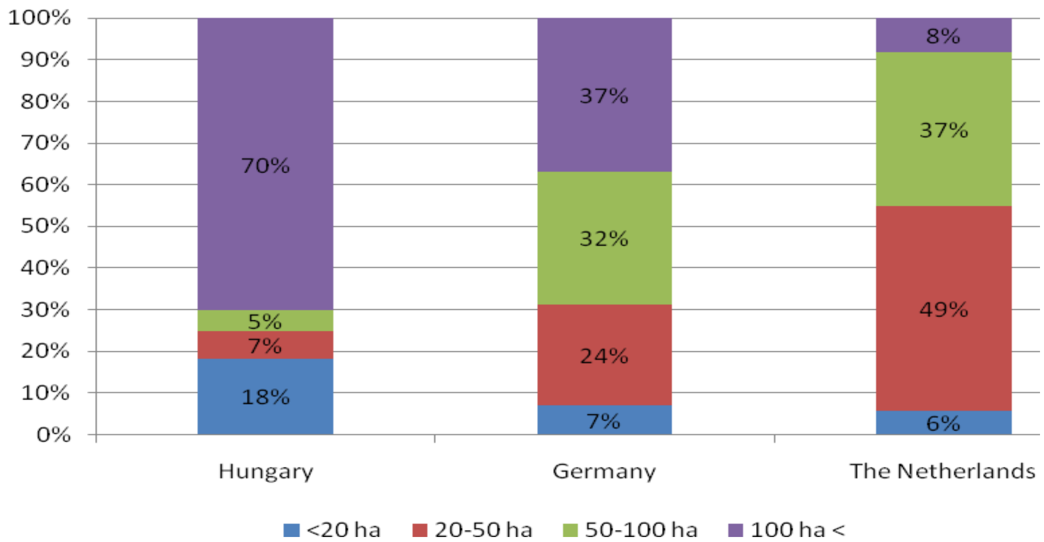



Source: FAOSTAT 2010

*Figure 2.5: Average cow milk price tendencies from 1991 to 2005*

## 2.2 Introducing the Dutch, German and the Hungarian dairy sector

The European Union is the largest milk producer in the world and the EU dairy sector is one of Europe’s most important farming sectors. To compare the three countries dairy farms efficiency, it is essencial to examine the structural differences between the countries. Figure 2.6 presents the distribution of dairy farm livestock according their size in terms of agricultural area (ha) in percentage, which means how much land the dairy farms have in the different countries. The hungarian dairy farms are mainly large in terms of land. 70 percent of the farms use more than 100 hectares of land for their business. The German farm’s represent a mix of small (less than 50 hectares land), medium (between 50 and 100 hectares land) and big (more than 100 hectares) farms. The Dutch dairy sector consists of many small and middle-sized farms, with the big dairy farms accounting for only 8 percent of the whole land. The Hungarian dairy sector is land extensive in contrast to the Dutch dairy sector which is land intensive. This intensive farming practices can involve very large numbers of animals raised on limited land which require large amounts of food, water and medical inputs. The German dairy sector about the land use is somewhere in the middle of the other two examined countries. This specialisation will be discussed in later sections.



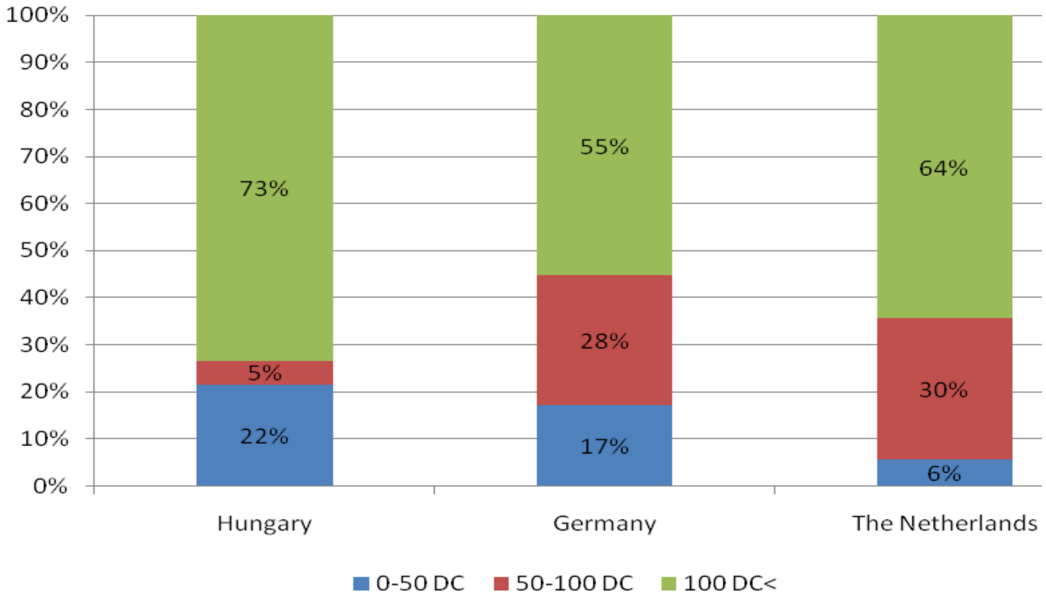

Source: EUROSTAT 2010.

*Figure 2.6: Distribution of dairy farms according their size in terms of agricultural area (ha) in percentage*

Another way to compare the dairy farms size examines the distribution of dairy farms according their size in terms of dairy cows (DC) in percentage (Figure 2.7). This figure presents the farms size regarding to the number of dairy cows instead of the agricultural land that the dairy farm use. Figure 2.7, shows that 73 percent of the Hungarian dairy livestock which means 0.19 million dairy cows live in big farms where there are more 100 dairy cows are kept. The average herd size is 22 dairy cows per holding (EUROSTAT 2010b).

The German farms characteristics are still the same as the previous comparison, so there are several types of farm working in Germany. 55% of the cows, which means 2.25 million dairy cows, live in big farms, where there are more than 100 dairy cows. The average size of the herd is 40.7 dairy cows per holding (EUROSTAT 2010c).

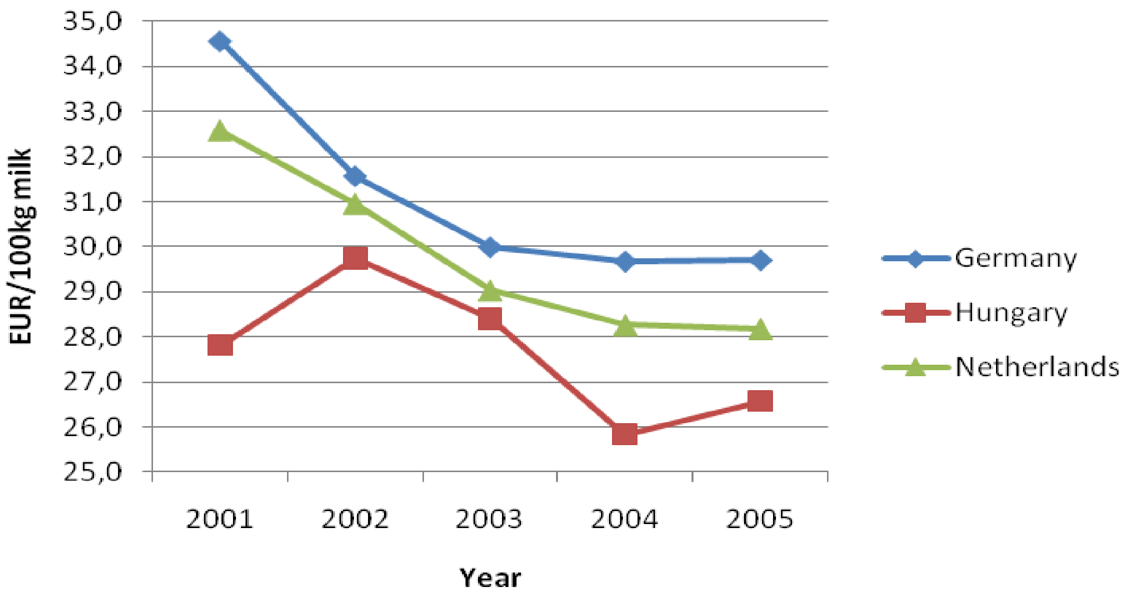
The Dutch farms are more specialised about dairy cows, so they own less land, but they keep the dairy cows in a big (more that 100 DC per farm) farms. 64 percent of the Dutch dairy cows, which means 0.946 million dairy cows live in dairy farms, with more than 100 cows. The average size of the herd is 59.9 dairy cows per holding (EUROSTAT 2010a).




Source: EUROSTAT 2010.

*Figure 2.7: Distribution of dairy farms according their size in terms of dairy cows (DC) in percentage*

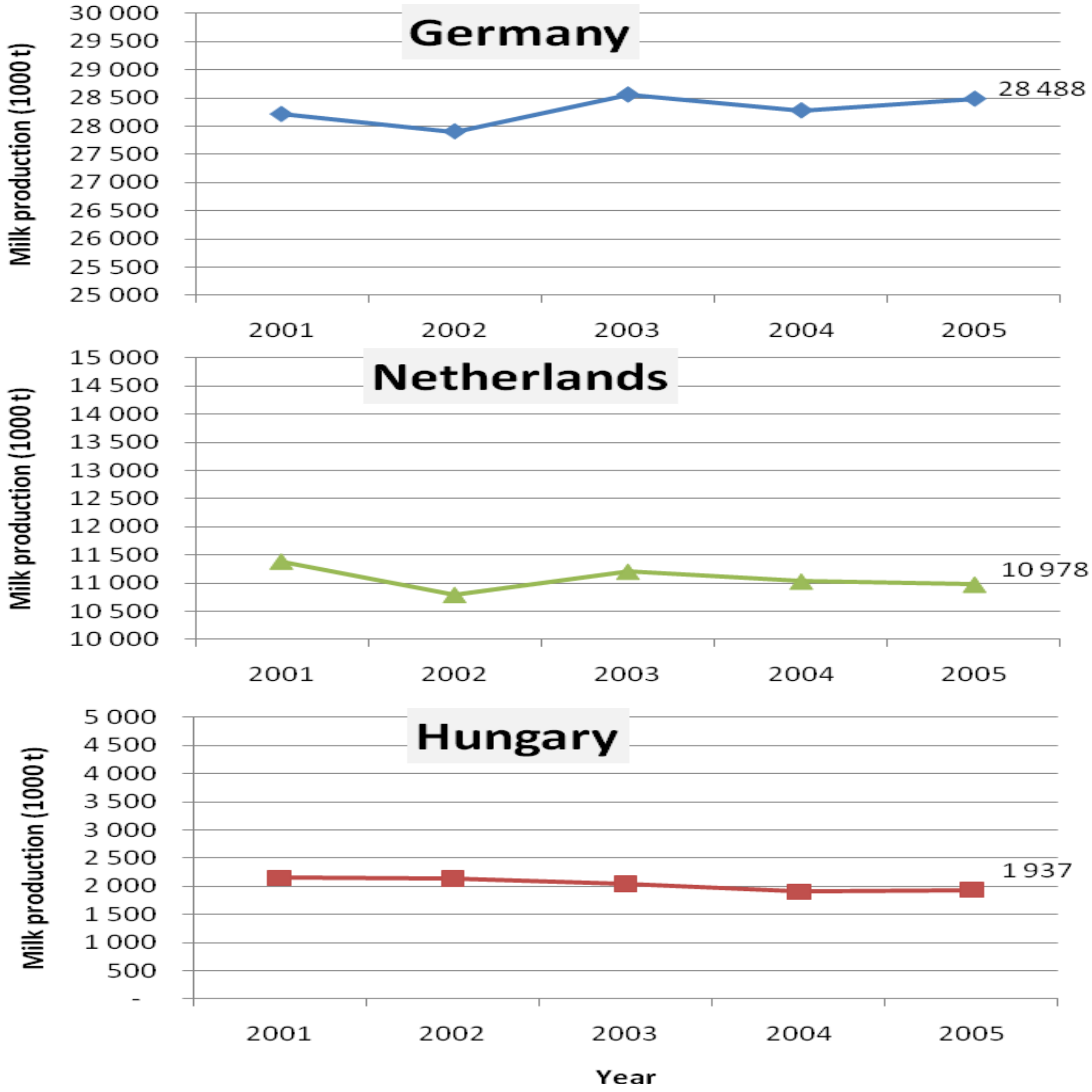
Regarding Figure 2.8, where the producer milk prices in the examined countries from 2001 to 2005 presented; the milk price showed a decreasing trend during the selected period. The German milk price is the highest (29.7 EUR/100kg milk) and exhibited the highest decrease from 2001, which is 14.08 percent. The Dutch milk price is 28.2 EUR/100kg milk, and presented 13.47 percent decrease from 2001. The Hungarian milk price level was always the lowest (26.6 EUR/100 kg milk) among the three countries, but it decreased only by 4.46 percent from 2001. Observable that after Hungary joined the European Union at 2004, the price level approached the European price level.




Source: EUROSTAT 2010.

*Figure 2.8: Producer milk prices in the examined countries from 2001 to 2005*

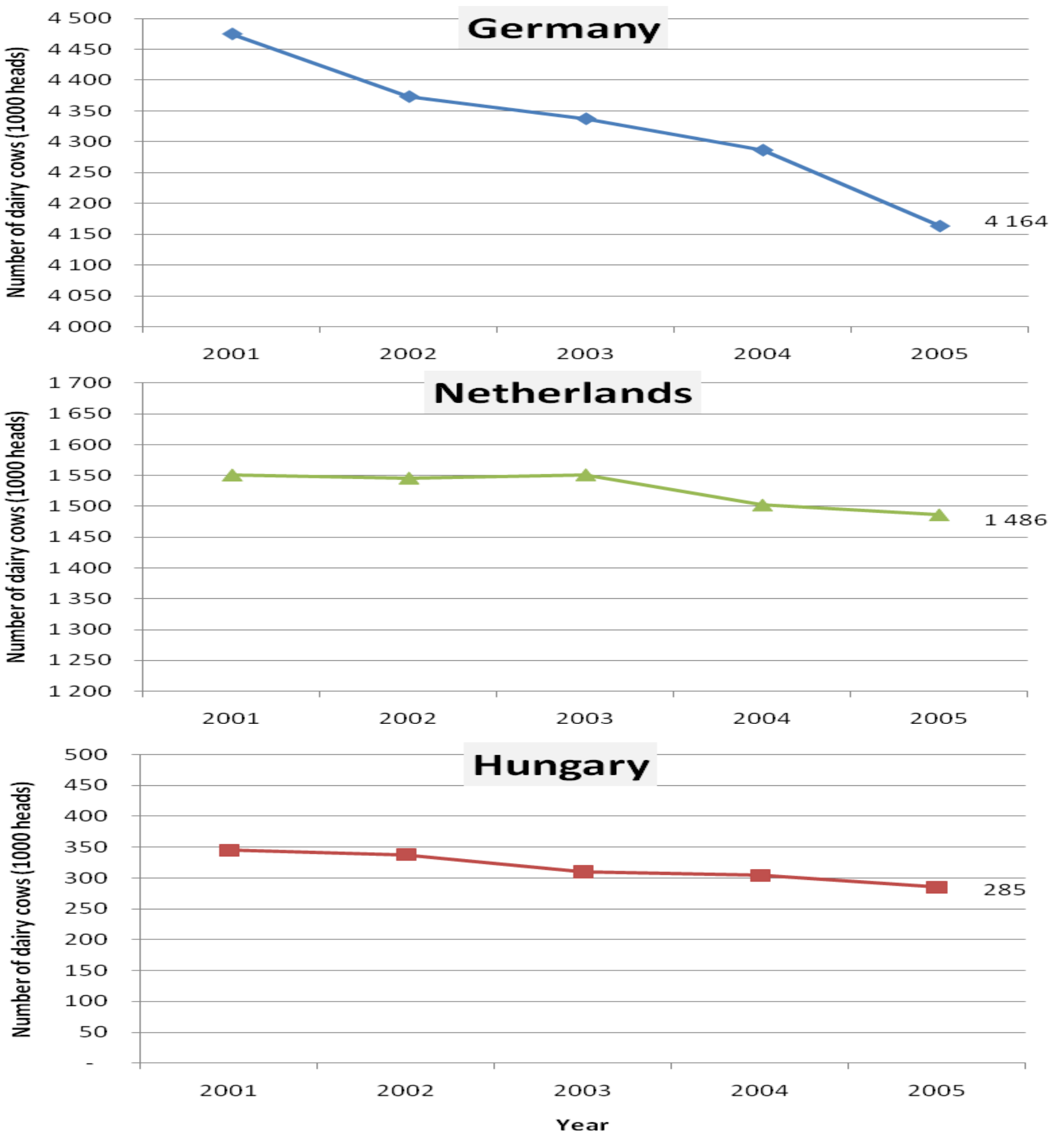
The EU-25 produced around 146 million tones of whole fresh cow milk in 2005 (EUROSTAT, 2010), which was 27.5 percent of the world production. The countries studied in this paper, namely Germany, Hungary and the Netherlands together represented around 27.74 percent of the total EU-25 production and 7.65 percent of the total world production (FAOSTAT 2010). According to Figure 2.9, which represents the milk production in the examined countries over the period of 2001 to 2005, the three countries milk production was relatively stable, as was the EU-25 production in this period. Among the three countries Germany is the largest milk-producing country with 28.49 million tonnes. The Netherlands and Hungary produced 10.98 million and 1.94 million tonnes respectively.




Source: EUROSTAT 2010.

*Figure 2.9: Milk production in the examined countries from 2001 to 2005*

The milk production was stable, but a small reduction was observed on the number of dairy cows (Figure 2.10) during the examined period. The country with the biggest cow population was Germany (4164 million heads in 2005) and the reduction was approximately 7 percent from 2001 to 2005. The Dutch dairy cows' number was 1486 million heads in 2005, which was quiet stable during the examined period. However a 4.2 percentage decrease occurred from 2001 to 2005. The Hungarian dairy cows' number was 285 thousand heads in 2005, which decreased 17.4 percent from 2001. Hence this was the highest decrease among the three countries.




Source: EUROSTAT 2010.

*Figure 2.10: Number of dairy cows in the examined countries from 2001 to 2005*

An interesting observation is that during the examined period the number of cows decreased in all chosen countries, but the milk production was quite stable. This was caused by the increasing performance of the cows. The average milk production per year per cow (Table 2.1) is the highest in The Netherlands (7615 kg); and lower in Germany (6984 kg) and Hungary (6850 kg).

The milk production per operating cost indicator calculated by the average milk production per farm divided by the livestock-specific operating cost (feeding cost, herd renewal purchases, milk levy and other specific costs) and the non-specific cost (machinery and building upkeep, energy cost, contract work, taxes and other dues, other direct inputs cost). This indicator represents the partial operating technical productivity, which is the highest in the Netherlands and lowest in Germany.

The next indicator is the milk production per total labour index, which shows the labour productivity among the three countries. This indicator is also the highest in the Netherlands, but the lowest in Hungary. About the labour use Hungary is use their labour extensively; on the other hand the Netherlands and Germany use intensively (Table 2.1).

The milk production per forage area index presents the land intensity of the dairy farms, which is the highest in the Netherlands and lower in Germany and Hungary. The Hungarian result is really low, 46% of the Dutch index, which shows that the Netherlands use extremely high land intensive technology.

The milk production per total input index shows the milk production related with the input costs (operating cost and fixed cost), where the highest result came from Hungary and the lowest from Germany. That index presents the ratio of the milk production and the total inputs. The last index called milk production per farm shows that the Hungarian farms use less input and produce more output in the farm level, because of the size effect, thus we conclude that the Hungarian farms are larger than the Dutch or the Germans.

Based on Table 2.1 the Dutch farms are more efficient regarding the technical partial productivity indexes. It seems that after the dairy quota system abolishment the Dutch farmers will increase their production potential and they will reach the best efficiency results among the three countries. After the quota system abolishment the Hungarian farms should have to increase their technical efficiency, otherwise they will decrease their production potential, now it seems that they are producing extensively, but in a big volume per farm. The German farms are lied in between of the other two countries.

*Table 2.1: Partial productivity indicators in the examined countries in 2005*

| | Germany | Hungary | The Netherlands |
|---|---|---|---|
| Milk production per cow (kg/DC) | 6 984 | 6 850 | *7 615 |
| Milk production per total operating cost (kg/€) | 1 828 | 2 900 | *3 369 |
| Milk production per total labor (kg/AWU) | 172 464 | 85 374 | *333 553 |
| Milk production per forage area (kg/ha) | 7 324 | 5 849 | *12 572 |
| Milk production per total input (kg/€) | 939 | *1 928 | 1 603 |
| Milk production per farm (kg/farm) | 332 856 | *584 814 | 540 356 |

AWU: annual working unit; DC: dairy cow; *the best result among the three countries

Source: FADN REPORT 2010.

So far we measured the efficiency only through partial productivity indicators. Although it is impossible to decide which counties technical efficiency is the highest. So far the different countries measuring was limited by measuring one input and one output performance of the farms. Thus the measuring of the inputs and the outputs was separately, during the following chapters the efficiency performance measuring regard with respect to all inputs and all output as many authors called (Farrel, 1957; Begum el. al. 2009; Coelli et. all 2005., Tauer, 1998; Jaforullah and Whiteman, 1999; Stokes et al., 2007; Kumbhakar and Lovell , 2000; Emvalomatis, 2010) in the literature the *"multiple input and output measurement"*.

## 2.3 Measuring efficiency

Measuring the productive efficiency of the dairy sector is important to both the practical experts and the economic policy makers. *"If economic planning is to concern itself with practical industries, it is important to know how far a given industry can be expected to increase its output by simply increasing its efficiency, without absorbing further resources."* (Farrel, 1957) This chapter introduces the mathematical definition of efficiency and two methods which can measure technical efficiency; namely data envelopment analysis and stochastic frontier analysis. The last section gives a comparison of the two methods and presents the main assumptions and differences of the methods.

### 2.3.1 Mathematical definition of efficiency

Measuring efficiency is a widely used concept in economics. Economic (or overall) efficiency expressed as a combination of technical and allocative (or price) efficiencies. Technical efficiency is the ability of the farmer to obtain maximal output from a given set of inputs while allocative efficiency measures the ability of the farmer to use inputs in optimal proportions, given their input prices and technology (Begum el. al. 2009; Coelli et. all 2005). There have been several methods to measuring efficiency; the generally used methods are data envelopment analysis (DEA) and stochastic frontier analysis (SFA), which involve mathematical programming and econometric methods, respectively.

Farell (1957) distinguishes input and output orientated measures depending on which factor we assume altering. So in the input orientated measure the input quantities changing without changing the output quantities. The assumed objective is to reduce the input quantities as much as possible, without changing the output quantities.

Illustrating the input-orientated measure according to Farell (1957) and Coelli et. al. (2005) example, considering a firm with a single output (y), under the assumption of constant returns to scale (which allows the technology to be represented using the unit isoquant) and using two inputs ($x_1$ and $x_2$) in Figure 2.11. The fully efficient (which is a theoretical frontier) firm represented by SS'. The firm defined by point P is inefficient. The technical efficiency of the firm is the ratio of 0Q and OP. The efficiency score will lie between zero and one, and indicate the degree of technical inefficiency of the actual (P) firm. Firm Q is technically efficient, because it is on the SS' isoquant: there is no reduction of inputs possible without reduction in output (Figure 2.11).

For instance, if a firm's technical efficiency is 80% (the technical efficiency score is 0.8), that means 0Q line is 80 percent of the whole 0P line, thus the QP line is 20 percent of the whole 0P line, so firm P is 20 percent inefficient in that case comparing with firm Q (which is a fully efficient firm). Thus if we assume input-orientated technical efficiency of 80 percent for a farm, that means the farm can reduce inputs by 20 percent without changing output.

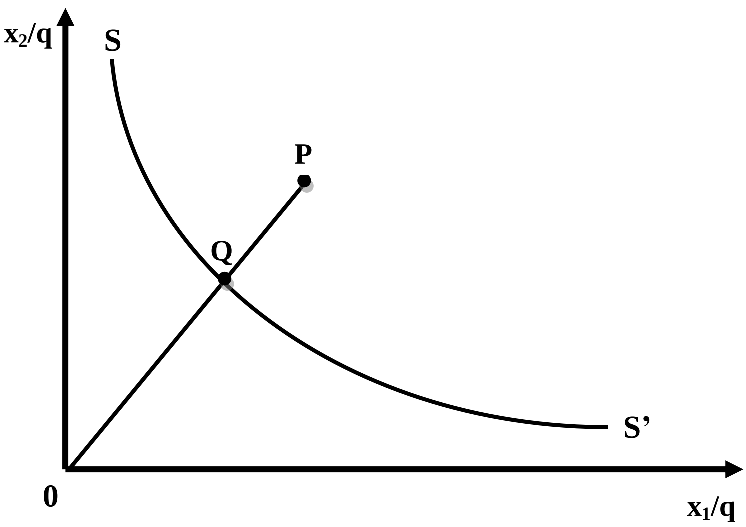



*Figure 2.11: Technical efficiency from an input orientation*

The other measure of efficiency regarding to Farell (1957) and Coelli et. al. (2005) is the output orientated measure, which is the opposite of the input orientated. By this measuring the question is: *"By how much can output quantities be proportionally expanded without altering the input quantities used?" (Coelli et. al. 2005).* If the technology is characterized by constant returns to scale the two orientations produce the same technical efficiency score. Differences, however, appear under variable returns to scale.

Figure 2.12 presents the technical efficiencies from an output orientation, here following Coelli et. al. (2005) and considering a firm with two outputs ($q_1$ and $q_2$) and a single input ($x_1$) and keep the input quantity fixed (because it is an output orientation measure), ZZ' represents the production possibility curve and point A the inefficient firm (Figure 2.12).

The distance AB measure the technical inefficiency, hence the output orientated technical efficiency is the ratio of 0A and 0B, which shows the percentage by which outputs could be increased without requiring extra input.

The input and the output orientated models estimate the same frontier and identify the same set of firms as being efficient, the difference is the efficiency measures associated with the inefficient firms that may differ between the two methods (Coelli et. al. 2005).

In practise the efficient isoquant is not known, the researchers have to estimate it from the sample data using different kinds of analyses. These will be introduced in the following sections. These analyses are the non-parametric data envelopment analysis and the parametric stochastic frontier analysis.

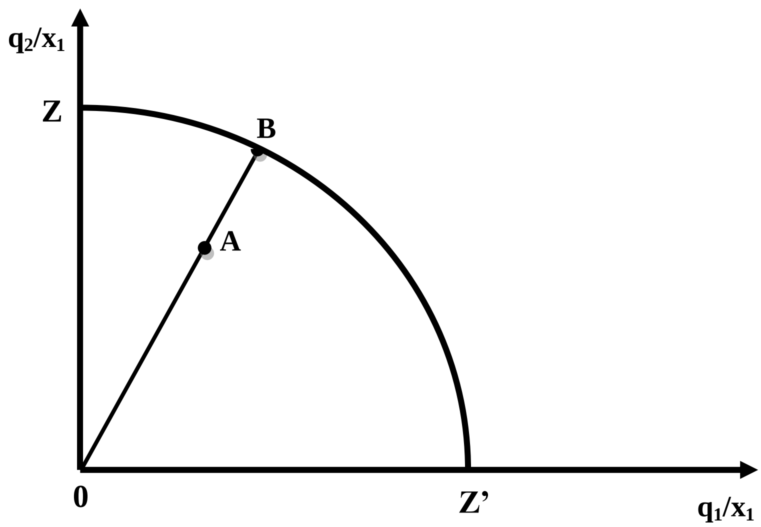



*Figure 2.12: Technical efficiency from an output orientation*

### 2.3.2 Introducing the Data Envelopment Analysis (DEA) Method

This section is divided into several subsections. The first part introduces a basic DEA model, in which a constant returns to scale (CRS) technology is assumed, the following part describes a more general variable returns to scale (VRS) DEA model.

The framework for the Data Envelopment Analysis (DEA) approach has been introduced by Farrell (1957) at first and popularized by Charnes, Cooper & Rhodes (1978). Data envelopment analysis is a non-parametric mathematical programming approach to frontier estimation.

The first and widely applied model was the input orientated CRS models, which solves the following linear programming problem for each firm to obtain the efficiency score:

$$\max_{u,v} (u'y_i / v'x_i),$$
$$\text{constrains:} \quad u'y_j / v'x_j \leq 1, \qquad j=1,2,...,N, \qquad (1)$$
$$u,v \geq 0$$

Where regarding to Coelli et. al. (2005), assuming K inputs and M outputs for each N firms. For the i-th firms the column vectors are represented by $x_i$ and $y_i$ respectively. X indicate the K*M input matrix and Y shows the M*N output matrix for all N firms. To measure efficiency we want to obtain the measure of the ratio of all outputs over all inputs, like $u'y_i / v'x_i$ where u represents the M*1 vector of output weights and v represents the K*1 vector of input weights. The obtained efficiency score will be less than or equal to one. There is one problem with tis formulation, because it has an infinite number of solutions. Charnes, Cooper & Rhodes (1978) solve it by adding one constrain $v'x_i = 1$ and reformulate the objective function a bit, this form we known as the multiplier form of the DEA. Using the duality linear programming method from the multiplier formula the envelopment form can get, which is the following:

$$\min_{\theta\lambda} \theta,$$
$$\text{constrains:} \quad -y_j + Y\lambda \geq 0, \qquad (2)$$
$$\theta x_i - X\lambda \geq 0,$$
$$\lambda \geq 0,$$

where λ represents the vector of peer weights. θ is a scalar and the value of it will be the efficiency score for the i-th firm, the value of 1 indicate the frontier and hence a technically efficient firm (but in practise it is not exist). This linear programming problem must be solved N times, once for each firm in the sample. Hence, each firm has its own θ efficiency score (Coelli et. al. 2005).The points of the fully efficient firms determine the fully efficient frontier line.

Regarding to the Eq. (2), takes the *i-th* firm and then seeks to radially contract the input vector, $x_i$, as much as possible, while still remaining within the feasible input set. The inner boundary of this set is a piece-wise linear isoquant (refer Eq. (1)), determined by the observed data points which are the firms in the sample. The radial contraction of the input vector, $x_i$, produces a projected point, (Yλ, Xλ*)*, on the surface of this method. This projected point is a linear combination of these observed data points. The constraints in Eq. (2) ensure that this projected point cannot lie outside the feasible set (Coelli et. al. 2005).

The constant returns to scale assumption is acceptable if the firms in the sample are operating at an optimal scale, but in practise the firms with imperfect competition do not behave like that. Banker, Charnes and Cooper (1984) suggested a model which can deal with variable returns to scale (VRS) situation. This model is quite similar to the CRS model except by adding a convexity constraint (N1'λ = 1) to the model, which accounts for the variable returns to scale. The model regarding to Banker, Charnes and Cooper (1984) and Coelli and Perelman (1996) presents an output oriented model, when the firms have fixed quantity of resources (capital, labour, livestock, land) and want to produce output (milk, calf) as much as possible. This model is very similar to the input orientated model. So the formula of an output orientated VRS model is the following:

$$\max_{\phi\lambda} \phi,$$

$$\text{constrains:} \quad -\phi y_j + Y\lambda \geq 0, \quad x_i - X\lambda \geq 0, \quad N1'\lambda = 1, \quad \lambda \geq 0, \qquad (3)$$

where the N1 is an N*1 vector of ones moreover $1 \leq \phi < \infty$ and $\phi$-1 is the proportional increase in output that could be achieved by the i-th firm, with input quantities held constant. 1/ $\phi$ determine the technical efficiency score, which lies between zero and one.

The DEA VRS formula envelopes the data points more tightly and provides higher or equal efficiency scores than the CRS model. The difference between the VRS and CRS technical efficiency scores is the scale inefficiency.

### 2.3.3 Introducing the Stochastic Frontier Analysis (SFA) Method

DEA estimates the unknown production frontier non-parametrically as the previous section introduced it, but there is another method, called stochastic frontier analyses, which uses parametric method for this estimation. The basic deterministic production frontier model, which does not account for errors and other statistical noise is due to Aigner and Chu (1968). Thus all deviations from the frontier are assumed to be the result of technical inefficiency. Independently Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) proposed a new stochastic frontier production function model which accounts for the measurement errors and other statistical noise; is the following:

$$q_i = \exp(\ f(x)\ ) * \exp(v_i) * \exp(-u_i) \qquad (4)$$

where the i-th firm produces one output ($q_i$); the deterministic component of the model is: exp( f(x) ), where f(x) is the production function and the inefficiency part is: exp($-u_i$) . This model takes into account the statistical noise as well, which is captured by the middle (exp($v_i$)) part of the model.

So the output oriented measure of technical efficiency is the ratio of observed output to the curresponding stochastic frontier output (Coelli et. al. 2005):

$$TE_i = \frac{q_i}{\exp(f(x)+v_i)} = \frac{\exp(f(x))+v_i-u_i)}{\exp(f(x)+v_i)} = \exp(-u_i) \Rightarrow -\log TE_i = u_i \qquad (5)$$

So the *i*-th firm technical efficiency ($TE_i$) is the exponential value of the negative inefficiency value ($-u_i$) of the *i*-th firm, or the negative logarithm of the *i*-th firm technical efficiency ($TE_i$) gives the *i*-th firm inefficiency value ($-u_i$). These technical efficiency value lies between zero and one and measures the output of the *i-th* firm relative to the output that could be produced by a fully-efficient firm using the same input vector.

Extening these models to estimate the production frontier with multiple output, the distance function can be used. It is necessary to use output distance functions, if the firms have more control over outputs than inputs.

The definition of the output distance function is:

$$D_0=(x, y) = \min \{ \theta : \frac{y}{\theta} \in P(x)\} \qquad (6)$$

where the x is the vector of inputs and y is a vector of outputs, P (x) is the set of output vectors that can be produced by x, θ represents the technical efficiency score, which lies between zero and one. An output distance function gives the maximum linear expansion of an output vector so that this vector reaches the boundary of the output possibilities set for given input (x). The output distance function assumes values in the interval zero and one and the locus of points for which $D_O$ (x, y) = 1 defines the boundary of the output possibilities set. From the definition the output distance function is linearly homogeneous in y.

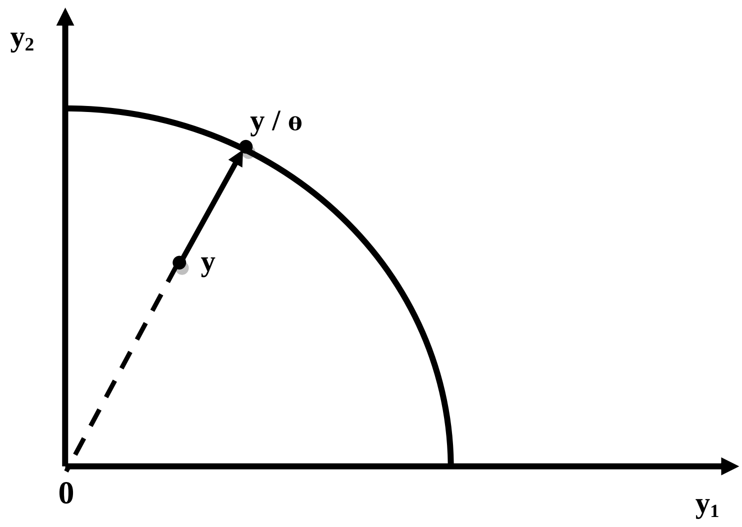



*Figure 2.13: Output distance function*

Figure 2.13 shows the graphical representation of an output distance function for a given set of inputs (x). The curve represents the boundary of the production possibilities set (P(x)). For any observed y, the distance function returns a number ɵ ϵ (0, 1) such that y/ɵ is on the boundary of P(x).

By definition the output distance function is homogeneous of degree 1 in outputs. Using this property, we can get an estimable from. So the multiple output distance function consider multiple output with the given set of inputs, thus the formula is the following if there is multiple inputs $M^{th}$ output for N firms:

$$-\log y_M = \log D_0 (x_i, y_i, t) - \log TE_i - v_i \qquad (7)$$

where $y_M$ is the $M^{th}$ output; $x_i$ the inputs, $y_i$ the outputs, t is the time variable, $TE_i$ is the technical efficiency like the ɵ previously and $v_i$ is the noise.

### 2.3.4 Comparison of DEA and SFA

The Data Envelopment Analysis (DEA) is an optimization technique built to measure the relative efficiency of a farm in the presence of multiple inputs and products (outputs). DEA provides a method to compare efficiency without knowing the production function, namely, without needing to know a functional relationship between inputs and outputs (Hormazábal and Wyngard 2007).

A problem with the DEA frontier is that it does not account for measurement errors and other sources of statistical noise. Thus all deviations from the frontier are assumed to be the result of technical inefficiency (Coelli et. al. 2005). But the DEA doesn’t need functional specification, thus there is less room for misspecification. The strength of DEA is that it does not require any assumptions about the functional form. The major weakness of DEA is that it is deterministic (Begum et al. 2009).

The stochastic production frontier approach proposed by Aigner et al. (1977) has been used by many researchers for the efficiency of business services sectors. The basic model assumes that total production deviates from the optimal production by a random noise and an inefficiency component. It is assumed to follow a symmetric normal distribution around the frontier that captures random disturbances, a phenomenon beyond the control of management. The inefficiency component can be distributed as half-normally (or for instance: truncated normally, exponentially or gamma) and represents the individual firm's deviations from the efficient frontier due to factors under management control. The stochastic frontier approach requires the specification of a general form of production function like Translog or Cobb-Douglas functional form. (Kasman and Turgutlu, 2007).

The stochastic frontier analysis (SFA) objective is the same as the DEA: to measure efficiency. Referring to Coelli et. al. 2005, the main advantages of this analysis over the DEA are: it accounts for noise and it can be used to conduct conventional tests of hypotheses. But the SFA models is usually more complicated than the DEA models, so the main disadvantages of the SFA with regard to the model specification. For example the distributional form for the inefficiency term and the production function.

Disadvantage of the SFA and DEA methods, that are relatively complicate to calculate it, comparing with the least squares econometric methods and the total factor productivity (index numbers) methods (Coelli et. al. 2005). Applying the frontier approach requires a number of firms in each time period and the outliers may influence of the results.

If we use DEA or SFA methods, the efficiency scores will only relative to the best firms in the sample, thus the inclusion of extra firms or other region or country may reduce efficiency scores. It can happen if we assume a common frontier among the countries. These models are really sensitive for the outliers as well, which may cause bias results, like measurement error and other noise may influence the shape and position of the frontier (Coelli et. al. 2005).

# 3 Materials and methods

This chapter firstly introduces the FADN database which has been used for this thesis. It includes yearly data from 2001 to 2005 for different dairy farms in Germany, the Netherlands and from 2001 to 2008 for Hungary. The second part presents the empirical specification of the Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) models.

## 3.1 Description of the data

In this research we use a database from the European Farm Accountancy Data Network (FADN). The concept of the FADN was launched in 1965, when Council Regulation 79/65 established the legal basis for the organisation of the network. It consists of an annual survey carried out by the Member States of the European Union (EU). The agencies responsible in the Union for the operation of the FADN collect every year accountancy data from a sample of the agricultural holdings in the European Union. Derived from national surveys, the FADN is the only source of micro-economic data that is harmonised; because the bookkeeping principles are the same in all countries. Holdings are selected to take part in the survey on the basis of sampling plans established at the level of each region in the EU. The survey does not cover all the agricultural holdings in the EU, but only those which due to their size could be considered commercial. The methodology applied aims to provide representative data along three dimensions: region, economic size and type of farming (FADN 2010a).

Currently, the annual sample covers approximately 80.000 holdings. They represent a population of about 5.000.000 farms in the 25 Member States, which cover approximately 90% of the total utilized agricultural area (UAA) and account for about 90% of the total agricultural production of the EU. It is expected that for the EU-27, that is including Bulgaria and Romania, the FADN would represent about 6.400.000 farms. The information collected, for each sample farm, concerns approximately 1000 variables (FADN 2010b).

To ensure that this sample reflects the heterogeneity of farming before the sample of farms, Liaison Agencies stratify the field of observation is defined according to 3 criteria: region, economic size and type of farming. Farms are selected in the sample according to a selection plan that guarantees its representativity. An individual weight is applied to each farm in the sample, this corresponding to the number of farms in the 3-way stratification cell of the field of observations divided by the number of farms in the corresponding cell in the sample. This weighting system is used in the calculation of standard results. The database contains farm level data, where the input and output data express with monetary units (€). The dataset organised by yearly for every farm, so this makes the panel dataset (FADN 2010c).

In this research we selected the dairy farms from Germany, Hungary and the Netherlands from 2001 to 2005. We focussed mainly on those dairy farms, whose revenues from cow's milk production are at least 75% of their total revenues for every year.

We use two outputs in our model, the revenues from cow's milk production and the revenues from other outputs. This other output revenues includes revenues from beef and veal and other output production that a dairy farm can produce. For the better estimation to account for the dependence of revenues on inflation, the output revenues and the input costs are deflated with country-wide price indices for each category of products, with prices obtained from EUROSTAT.

The analysis uses six deflated (base year is 2000) inputs categories, which cover the whole input side of the dairy business. These categories are the following:

1; Capital (K) consists of the buildings and fixed equipment like: tractors, lorries, milking machines, cleaning machines, feeding automats.

2; Labour (L) is measured in working hours and includes both family and hired labours.

3; Land (A) is measured in hectares, and includes the total utilized agricultural area (UAA) of the holding. Does not include areas used for woodland, roads, non-farmed areas.

4; Total material inputs (M) includes all deflated farm specific costs, that arise in the dairy business like: seeds and plants, fertilizers, crop protection, crop and livestock-specific cost (storage cost, marketing cost, veterinary cost) and energy (fuel, electricity, heating) costs.

5; Livestock (S) is measured in standardized livestock unit (LSU) which is the total number of livestock heads on the farm aggregated with European standard weight coefficients. In our case the LSU includes female bovine animals, which have calved and are held principally for milk production for human consumption and other cattle. The weights for dairy cows are 1, while the younger than two years cattle weights are 0.4 to 0.6.

6; Purchased feed (F) is measured in deflated monetary value, and includes purchased feed and concentrates for grazing and home-grown livestock, but excludes the value of feed produced within the farm.

The following table contains the descriptive statistic from the used dataset:

*Table 3.1: Variable averages and standard deviations (SD) in the examined countries*

| | Germany | | The Netherlands | | Hungary* | |
|---|---|---|---|---|---|---|
| | Average | *SD* | Average | *SD* | Average | *SD* |
| Milk revenues (€) | 104 587 | *122 106* | 186 221 | *105 997* | 154 573 | *364 781* |
| Other revenues (€) | 32 553 | *39 187* | 32 807 | *25 902* | 52 265 | *140 798* |
| Capital (€) | 167 258 | *162 329* | 196 327 | *145 140* | 89 124 | *144 576* |
| Labor (AWU) | 4 085 | *4 245* | 4 251 | *1 753* | 16 038 | *32 601* |
| Land (UAA) | 63 | *73* | 50 | *29* | 164 | *339* |
| Material inputs (€) | 44 699 | *52 518* | 52 230 | *26 455* | 81 718 | *223 520* |
| Livestock (DC) | 92 | *91* | 113 | *61* | 159 | *326* |
| Purchased feed (€) | 20 448 | *33 505* | 33 099 | *22 308* | 58 596 | *148 720* |

AWU: annual working unit; UAA: utilized agricultural area; *time interval is 2001 to 2008 for Hungary

Source: Own calculation based of the FADN database 2001-2005.

Table 3.1 prove the same results, as we have seen in the Section 2.2, where the three countries dairy sector have been introduced. Here the selected farms represent their countries quiet well. In the Netherlands, we can see the highest milk revenue per farm, Hungary is in the second place, but the standard deviation value is three times higher than the other countries, so this average doesn't make a good representation of the whole sample.

The input side of the dataset prove the previous sections statement, which is for instance the Hungarian dairy farms are labor-extensive; on the other hand the Netherlands and Germany use intensively or an other statement wes that the Hungarian dairy sector is land extensive in contrast to the Dutch dairy sector which is land intensive.

## 3.2 Empirical specification of the DEA model

The objective of this section is to compare dairy farms efficiency in Germany, Hungary and the Netherlands. We assume the farms produce two kinds of output, which are the revenues of cow's milk and the revenues of other output. The other output includes all the other outputs of the farms which are valuable except the cow milk. The other outputs variable is equal to the difference of the total outputs and the sum of the farm use production and the total cow milk production.

The model uses output orientation variable returns to scale (VRS) configurations. We assume output orientation, instead of the input orientation, however the European dairy market is restricted by the quota system, but the dairy farmers can trade with the quotas among themselves (inside the country) in every examined country. On the other hand we assumed that the dairy farms easily can buy more quotas if it's necessary than change the quantity of their inputs. So we assumed that the inputs are more fixed than the outputs. This output orientated VRS model is quite similar to the constant returns to scale (CRS) model except by adding convexity constraint (N1'λ = 1) to the model, which account for the variable returns to scale.

The model is an output oriented model, when the firms have fixed quantity of resources (capital labor, land, total material inputs, livestock and feed) and wanted to reach as much revenues, which is measures the quantity of the outputs (milk, other) as possible. The farms have six inputs, which covers approximately all the input what they use to produce milk. These inputs are: capital (machinery and buildings), labor, land, total material inputs, livestock and feed.

We estimate the technical efficiency for the three countries for every year individually. That procedure gives the German and the Dutch technical efficiency score from 2001 to 2005 and from 2001 to 2008 for Hungary, thus we can create an average technical efficiency score for the countries, to compare them.

## 3.3 Empirical specification of the SFA model

In our model the dairy farms produce two outputs, milk and other output, which includes beef and veal, manure and other outputs. This multiple output technology better represented by a distance function rather, than a single production function. This model uses output distance function; because we assume that the farmers try to increase the quantity of outputs from the given quantity of inputs. In the stochastic frontier analysis (SFA), which is a parametric method, this distance function is specified as translog function in inputs (x), outputs (y) and time (t):

$$
\begin{aligned}
\log Do(x_i, y_i, t) = & \alpha + \sum_k \beta_k \log(x_{ki}) + \sum_k \gamma_k \log(y_{ki}) \\
& + \frac{1}{2}\sum_k \sum_l \delta_{kl} \log(x_{ki}) \log(x_{li}) \\
& + \frac{1}{2}\sum_k \sum_l \zeta_{kl} \log(y_{ki}) \log(y_{li}) \\
& + \frac{1}{2}\sum_k \sum_l \eta_{kl} \log(x_{ki}) \log(y_{li}) \\
& + \theta_1 t + \theta_2 t^2 \\
& + \sum_k \varepsilon_k t \log(x_{ki}) + \sum_k \psi_k t \log(y_{ki})
\end{aligned}
\tag{8}
$$

This output distance function (8) has different curvature in the input and output dimensions as well. To capture the effect of technological changes, we introduce the interaction terms as well. So finally the translog function makes every combination of the variables what we have in our models, which are the two outputs, the six inputs and the time.

Finally we have to normalise the model with one output, for instance we can choose the cow milk production as the normalizing output to get the following equation:

$$-\log y_{cmilk}=\log D_0 (x_i , y_{i/}y_{cmilk} , t) - \log TE_i + v_i \qquad (9)$$

where $y_{cmilk}$ is the cow milk output as a dependent variable; $x_i$ the inputs which are constants, $y_{i/}y_{cmilk}$ is the function of ($\log y_{others}$- $\log y_{cmilk}$) the outputs, t is the time variables, $TE_i$ is the technical efficiency and $v_i$ is the noise.

The data for all inputs and all outputs are normalized by their appropriate geometric means prior to estimation. That procedure makes the model's parameter estimates directly interpretable as distance elasticises evaluated at the geometric mean of the data.

In this thesis we use the Barttese and Coelli (1992) time-varying panel model to predict the technical efficiency on an individual firm at the particular time period. Our empirical example is the Dutch and German dairy farms data from 2001 to 2005 and for the Hungarian dairy farms from 2001 to 2008.

Barttese and Coelli (1992) considered a stochastic frontier production function with simple exponential specification of time-varying firm effects which incorporates unbalanced panel data associated with observations on a sample on N farms over T time periods. The model is the following:

$$Y_{it}=f(x_{it};\beta)\exp(V_{it}-U_{it}) \qquad (10)$$

and

$$U_{it}=U_i*\{\exp[-\eta(t-T)]\}, \quad i=1,2,....,N; \qquad (11)$$

where $Y_{it}$ represents the production for the i-th firm at the t-th period, $f(x_{it};\beta)$ the suitable function of a vector $x_{it}$, of factor inputs associated with the production of the i-th firm in the period t, vector $\beta$ is an unknown parameter; $V_{it}$ is assumed to be independent and identically distributed random errors; $U_i$ is assumed to be independent and identically distributed non-negative truncations of the normal distribution; $\eta$ is an unknown scalar parameter, T the set of the time periods, t is the time between the time period T.

# 4 Results

This section presents the results of the data envelopment analysis (DEA) and stochastic frontier analyse (SFA) for Germany, The Netherlands and Hungary. At the end of the section we will compare the results from the two analyses, and if there is any difference between them, we try to explain why can is happened.

## 4.1 Results of the DEA

Table 4.1 presents the estimated mean values of technical efficiency which on average for 2001 to 2005 is 83 percent assuming variable returns to scale (Vrste) for Germany with 982 observations per year. The scale efficiency is the ratio of the constant and variable returns to scale (0.80/0.83), which is on average 0.96 and indicate that the difference between the constant and variable returns to scale is only 4% which is close to constant returns to scale (CRS) part of the technology.

*Table 4.1: Summary of the technical efficiency in Germany*

| Year | Crste | Vrste | Scale | Observations |
|---|---|---|---|---|
| **2001** | 0.80 | 0.84 | 0.96 | 982 |
| **2002** | 0.80 | 0.83 | 0.96 | 982 |
| **2003** | 0.80 | 0.83 | 0.97 | 982 |
| **2004** | 0.80 | 0.83 | 0.97 | 982 |
| **2005** | 0.81 | 0.84 | 0.96 | 982 |
| **Average** | 0.80 | **0.83** | 0.96 | 982 |

Note: crste = technical efficiency from CRS DEA; vrste = technical efficiency from VRS DEA; scale = scale efficiency = crste/vrste
Source: Own calculation based of the FADN database 2001-2005.

*Table 4.2: Summary of the technical efficiency in The Netherlands*

| Year | Crste | Vrste | Scale | Observations |
|---|---|---|---|---|
| **2001** | 0.88 | 0.92 | 0.96 | 178 |
| **2002** | 0.89 | 0.92 | 0.97 | 178 |
| **2003** | 0.90 | 0.93 | 0.97 | 178 |
| **2004** | 0.89 | 0.92 | 0.97 | 178 |
| **2005** | 0.89 | 0.93 | 0.97 | 178 |
| **Average** | 0.89 | 0.92 | 0.96 | 178 |

Note: crste = technical efficiency from CRS DEA; vrste = technical efficiency from VRS DEA; scale = scale efficiency = crste/vrste
Source: Own calculation based of the FADN database 2001-2005.

Table 4.2 presents the estimated mean values of technical efficiency which on average for 2001 to 2005 is 92 percent assuming variable returns to scale (Vrste) for the Netherlands with 178 observations. The scale efficiency is on average 0.96 and indicates that the difference between the constant and variable returns to scale is only 4% which is also close CRS part of the technology.

The Hungarian database is quiet problematic, because of the small specialised dairy farm number of the sample. Thus to deal with the small sample size, we used aggregate dataset over time (from 2001 to 2008) and we weren't apply that restrictions, that the farms have to be in the database at least five years. Table 4.3 presents the estimated mean values of technical efficiency which on average for  2001 to 2008 is 90 percent assuming variable returns to scale (Vrste) for Hungary with 94,5 observation on average. The scale efficiency is on average 0.96 and indicates that the difference between the constant and variable returns to scale is only 4% like the other two countries. The high technical efficiency score caused by the low number of observation and the big specialised dairy farms in Hungary in the sample.

*Table 4.3: Summary of the technical efficiency in Hungary*

| **Year** | **Crste** | **Vrste** | **Scale** | **Observations** |
|---|---|---|---|---|
| **2001-2004** | 0.86 | 0.90 | 0.95 | 120.0 |
| **2005-2008** | 0.86 | 0.89 | 0.96 | 67.0 |
| **Average** | 0.86 | 0.90 | 0.96 | 94.5 |

Note: crste = technical efficiency from CRS DEA; vrste = technical efficiency from VRS DEA;  scale = scale efficiency = crste/vrste
Source: Own calculation based of the FADN database 2001-2008.

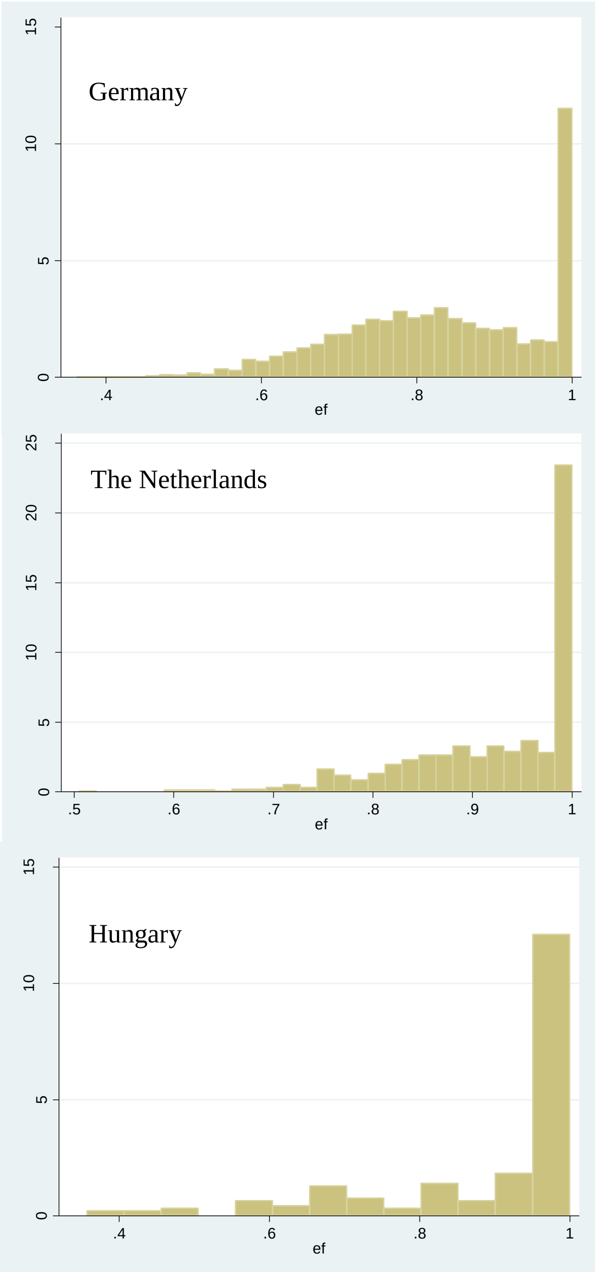
Table 4.4 presents the summary of the input slacks for the tree countries and shows that the Hungarian farms are less efficient regarding to the labor, land, livestock and feed inputs comparing the Dutch and the German farms. Hungary has to optimise these inputs to be more efficient. For instance if we assume input-orientation where the slack for the land variable is 15.8 hectare for a farm, that means the farm can reduce their land inputs by 15.8 hectare without changing the milk and the other outputs production, while moving along the frontier.

*Table 4.4: Summary of the input slacks in Germany, The Netherlands and Hungary*

| Year/input | Capital | Labor | Land | Material inputs | Livestock | Feed |
|---|---|---|---|---|---|---|
| Germany | 27681.00 | 323.20 | 9.89 | 949.57 | 3.03 | 478.77 |
| The Netherlands | 19455.85 | 369.13 | 3.06 | 2392.95 | 2.73 | 954.97 |
| Hungary* | 13186.57 | 1259.25 | 15.82 | 1060.04 | 5.43 | 1283.77 |

Note: * The Hungarian data are from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

Figure 4.1 presents histograms of the efficiency scores for the three examined countries. The shape of these graphs suggests a higher variability of efficiency score for Germany. The Hungarian graphs suggest less variability, but this is caused by the smaller number of observations. The Dutch left skewed distribution represents more efficient dairy farm comparing to the central skewed German distribution. The Hungarian graph is really fluctuating, so it is hard to interpret.




Note: * The Hungarian data are unbalanced from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

*Figure 4.1: Histograms of Efficiency Score Estimates (using DEA) for Germany, the Netherlands and Hungary*

Table 4.5 displays the average technical efficiency of the German, the Dutch and the Hungarian sample respectively, calculated under the assumption of a common frontier across three countries. In our case common frontier was necessary to eliminate the Hungarian low sample size and merge the three country dataset in one for the better estimation. The common frontier case assumes that the three counties can access to the most efficient technology. Results show that the superiority, in terms of technical efficiency average, of the Dutch sample in the dairy sector remains when using the common frontier (average technical efficiency (assuming VRS) of 0.9 for the Dutch farms, 0.77 for the German farms and 0.84 for the Hungarian farms). This suggests that more Dutch dairy farms are closer to the efficient common frontier than Hungarian or even the German dairy farms. Furthermore this suggests that, if it is assumed that there is common technology between Dutch, German and Hungarian farms, the Dutch farmers make a more efficient use of this technology in the dairy sector.

*Table 4.5: Comparing technical efficiency assuming commonl frontier using DEA*

| | **GERMANY** | | | | **The NETHERLANDS** | | | | **HUNGARY** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Year** | **Crste** | **Vrste** | **Scale** | **Obs** | **Crste** | **Vrste** | **Scale** | **Obs** | **Crste** | **Vrste** | **Scale** | **Obs** |
| **2001** | 0.74 | 0.76 | 0.98 | 982 | 0.87 | 0.89 | 0.98 | 178 | 0.70 | 0.82 | 0.87 | 51 |
| **2002** | 0.75 | 0.77 | 0.97 | 982 | 0.87 | 0.89 | 0.98 | 178 | 0.76 | 0.84 | 0.92 | 33 |
| **2003** | 0.74 | 0.76 | 0.97 | 982 | 0.88 | 0.90 | 0.98 | 178 | 0.75 | 0.80 | 0.93 | 20 |
| **2004** | 0.76 | 0.78 | 0.98 | 982 | 0.87 | 0.90 | 0.97 | 178 | 0.81 | 0.88 | 0.93 | 16 |
| **2005** | 0.75 | 0.78 | 0.97 | 982 | 0.88 | 0.91 | 0.97 | 178 | 0.79 | 0.84 | 0.94 | 15 |
| **Average** | 0.75 | 0.77 | 0.97 | 982 | 0.88 | 0.90 | 0.98 | 178 | 0.76 | 0.84 | 0.92 | 27 |

Note: crste = technical efficiency from CRS DEA; vrste = technical efficiency from VRS DEA; scale = scale efficiency = crste/vrste; Obs=Observation per year

Source: Own calculation based of the FADN database 2001-2005.

## 4.2 Results of the SFA

We know from SFA model specification section that the technical efficiency of the examined farm is defined by $TE_i = \exp(-u_i)$. This equation provides a basis for the prediction of the farm and the industrial (sectorial) technical efficiency. The industry efficiency is the average of the predicted efficiencies of the farms in the sample.

Table 4.6 reports the final results of the 3 countries parameter estimates of the first-order terms of the distance function. The full results table is in the Appendix section table A.4. All the estimated elasticties are statistically significant, except the labor parameter in the Netherlands and Hungary. It caused perhaps the lower sample size of these two countries.

*Table 4.6: Estimates of the Time-varying SFA model's parameters*

| | **Germany** | | | **The Netherlands** | | | **Hungary*** | | |
|---|---|---|---|---|---|---|---|---|---|
| **log_cmilk** | **Coef.** | **Std. Err.** | **p-value** | **Coef.** | **Std. Err.** | **p-value** | **Coef.** | **Std. Err.** | **p-value** |
| log_oth | 0.189 | 0.003 | 0.000 | 0.099 | 0.010 | 0.000 | 0.327 | 0.021 | 0.000 |
| log_K | -0.054 | 0.006 | 0.000 | -0.043 | 0.010 | 0.000 | -0.139 | 0.033 | 0.000 |
| log_L | -0.060 | 0.011 | 0.000 | -0.022 | 0.017 | 0.195 | 0.083 | 0.060 | 0.167 |
| log_A | -0.047 | 0.013 | 0.000 | -0.158 | 0.025 | 0.000 | -0.115 | 0.050 | 0.021 |
| log_M | -0.210 | 0.011 | 0.000 | -0.092 | 0.021 | 0.000 | -0.228 | 0.066 | 0.001 |
| log_S | -0.445 | 0.015 | 0.000 | -0.520 | 0.030 | 0.000 | -0.527 | 0.080 | 0.000 |
| log_F | -0.156 | 0.006 | 0.000 | -0.193 | 0.015 | 0.000 | -0.122 | 0.030 | 0.000 |
| trend | -0.016 | 0.002 | 0.000 | -0.020 | 0.003 | 0.000 | -0.004 | 0.010 | 0.693 |
| $\boldsymbol{\mu}$ | 0.207 | 0.033 | 0.000 | -0.070 | 0.223 | 0.753 | 0.389 | 0.635 | 0.540 |
| $\boldsymbol{\eta}$ | 0.001 | 0.006 | 0.850 | -0.044 | 0.015 | 0.003 | -0.019 | 0.077 | 0.808 |
| $\sigma^2 = \sigma_u^2 + \sigma_v^2$ | 0.064 | 0.008 | 0.000 | 0.079 | 0.040 | 0.000 | 0.056 | 0.006 | 0.000 |
| $\gamma = \sigma_u^2/\sigma^2$ | 0.882 | 0.015 | 0.000 | 0.953 | 0.023 | 0.000 | 0.154 | 0.204 | 0.000 |
| $\sigma_u^2$ | 0.057 | 0.008 | 0.000 | 0.076 | 0.040 | 0.000 | 0.009 | 0.012 | 0.000 |
| $\sigma_v^2$ | 0.008 | 0.000 | 0.000 | 0.004 | 0.000 | 0.000 | 0.047 | 0.011 | 0.000 |

Note: * The Hungarian data are unbalanced from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

The log_oth row results present the distance elasticties considering to outputs as measures of the curvature of the production possibilities frontier. That elasticity values mean, if the other output (which is the beef and veal and manure and other in our model) will increase 1 percent than cause 0.19% increase in the distance function, thus these farms will get closer to the production possibilities frontier in Germany. This elasticity value is 0.10% for the Netherlands and 0.33% for Hungary. The Hungarian elasticity value is the highest if we compare the three countries results respect to the other output, which means that the increase of the other parameter by 1 percent cause the highest increase in the distance function, thus this is the most sensitive countries for this parameter, which represents the beef and weal and other outputs of the dairy farming.

Considering the Hungarian other parameter's elasticity value, the elasticties implied by the linear homogeneity restrictions with respect to the cow milk output (log_cmilk) are about 0.67% for Hungary, which is the lowest marginal transformation rate of other output to milk. This number is 0.81% for Germany and 0.9% for the Netherlands.

The negative sign of the first-order terms in the Table 4.6 means that the increases in inputs push the production possibilities frontier outwards. Every input of the three countries has a negative elasticity of the distance function except the Hungarian labor parameter, but that parameter estimate is not significant statistically. For every countries the largest effect caused by the livestock input (log_S) for the outputs. The second important input for the outputs is the total material inputs (log_M) for Germany and for Hungary, but for the Netherlands the feed input (log_F) is that. The most interesting part is the third dominant input, which is the feed (log_F) for Germany; the land or area (log_A) for the Netherlands, and the capital (log_K) for Hungary. These third dominant inputs can give the varying characteristics of the three different countries dairy efficiency.

The negative trend parameter input means that every county has technological improvement over the years, which push the production possibility sets outwards over the years. Although the Hungarian technological improvement effect statically is not significant.

The scale elasticity of the distance function, which is calculated by adding the distance elasticties with respect to the six inputs are: -0.971 (p=0.02) for Germany; -1,027 (p=0.20) for the Netherlands and -1,047 (p=0.08) for Hungary thus we can assume that the examined countries dairies are operation in the increasing returns to scale part of the technology; except Germany, which dairies are operating the decreasing returns to scale part of the technology. That means for instance 1 percent increases for input side; generate 1.047% increase for the output side for Hungary; 1.027% for the Netherlands and 0.971% for Germany.

The estimate of η is positive for Germany, not suggesting improvements in technical efficiency over time. However, this effect is not statistically significant. For the Netherlands the η is negative, which is suggesting significant increasing in technical efficiency over these five years. For Hungary the η is negative, but not statistically significant.

STATA software parameterises the log-likelihood in terms of $\boldsymbol{\gamma = \sigma_u^2/\sigma^2}$. This estimate (0.953) is the highest for the Netherlands, meaning that much of the variation in the composite error term is due to the inefficiency component. The lowest γ is in Hungary (0.154) meaning that much of the variation in the composite error term is due to the statistical noise component and the less observation.

Table 4.7 presents the final results of the three countries technical efficiency score. The most efficient country comparing with their national production possibilities frontier is the Netherlands with 84%, the second is Germany with 76% and the third is Hungary with 68%. That means that the Hungarian dairy farms can improve their performance the most to reach their maximal reachable production level. The dairy farming technology is different for the three countries, that's why this comparison is more reliable than to assume a common production possibilities frontier for the three countries.

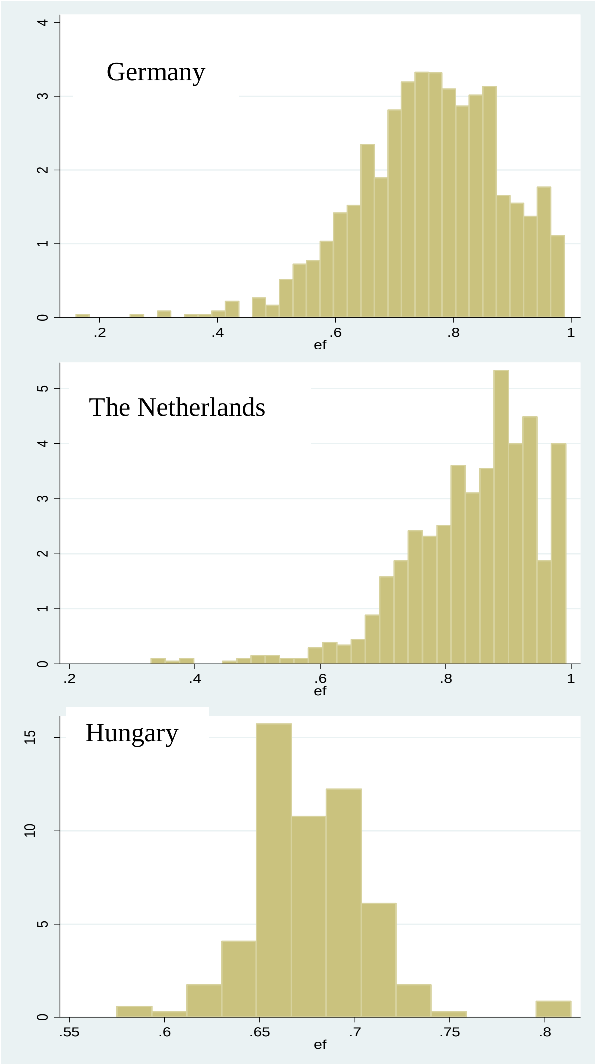
*Table 4.7: Comparing technical efficiency for the examined countries*

| Country | # of Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Germany** | *4910* | 0.76 | *0.12* | *0.16* | *0.99* |
| **The Netherlands** | *890* | 0.84 | *0.10* | *0.33* | *0.99* |
| **Hungary*** | *187* | 0.68 | *0.03* | *0.57* | *0.81* |

Note: * The Hungarian data's are unbalanced from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

Figure 4.2 presents histograms of the efficiency estimates for the examined countries. The shape of these graphs suggests a higher variability of efficiency score for Germany. The Hungarian graphs suggest less variability, but it caused the less number of observations. The Dutch left skewed distribution represents more efficient dairy farm comparing to the central skewed Hungarian distribution.

If we assume that every examined country can use the same technology, we should create a common production possibilities frontier for all the three countries. Table 4.8 contains summary statistics for the estimated efficiency scores in that case. The highest efficiency score is the Dutch efficiency with 84%, which is the same result as if we use national frontiers. We usually expected lower efficiency results assuming the common frontier. The German and the Hungarian efficiency score is much lower than the previous result, because of the really high Dutch performance.




Note: * The Hungarian data are unbalanced from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

*Figure 4.2: Histograms of Efficiency Score Estimates (using SFA) for Germany, the Netherlands and Hungary*

*Table 4.8: Comparing technical efficiency assuming common frontier using SFA*

| | **GERMANY** | | | **The NETHERLANDS** | | | **HUNGARY** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Year** | **Mean** | **SD** | **Obs** | **Mean** | **SD** | **Obs** | **Mean** | **SD** | **Obs** |
| **2001** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.63 | 0.15 | 51 |
| **2002** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.62 | 0.13 | 33 |
| **2003** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.63 | 0.10 | 20 |
| **2004** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.63 | 0.10 | 16 |
| **2005** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.55 | 0.17 | 15 |
| **Average** | 0.67 | 0.11 | 982 | 0.84 | 0.10 | 178 | 0.61 | 0.13 | 27 |

Note: SD=standard deviation; Obs=Observation per year
Source: Own calculation based of the FADN database 2001-2005.

*Table 4.9: Estimates of the Time-varying SFA model's parameters assuming common frontier*

| **log_cmilk** | **Coef.** | **Std. Err.** | **p-value** |
|---|---|---|---|
| log_oth | 0.178 | 0.003 | 0.000 |
| log_K | -0.059 | 0.005 | 0.000 |
| log_L | -0.053 | 0.010 | 0.000 |
| log_A | -0.023 | 0.012 | 0.058 |
| log_M | -0.221 | 0.010 | 0.000 |
| log_S | -0.523 | 0.014 | 0.000 |
| log_F | -0.165 | 0.006 | 0.000 |
| trend | -0.019 | 0.002 | 0.000 |
| $\mu$ | 0.365 | 0.028 | 0.000 |
| $\eta$ | -0.007 | 0.005 | 0.153 |
| $\sigma^2 = \sigma_u^2 + \sigma_v^2$ | 0.062 | 0.005 | |
| $\gamma = \sigma_u^2/\sigma^2$ | 0.868 | 0.012 | |
| $\sigma_u^2$ | 0.054 | 0.005 | |
| $\sigma_v^2$ | 0.008 | 0.000 | |

Source: Own calculation based of the FADN database 2001-2005.

In Table 4.9, similarly like in the separate national frontier cases the log_oth row results present the distance elasticties considering to outputs as measures of the curvature of the production possibilities frontier. That elasticity values means that if the other output (which is the beef and veal and manure and other in our model) will increase 1 percent than cause 0.18% increase in the distance function, thus these farms will get closer to the production possibilities frontier. Considering the common other parameter's elasticity value, the elasticties implied by the linear homogeneity restrictions with respect to the cow milk output (log_cmilk) are about 0.82% for the three examined countries.

The negative sign of the first-order terms in the Table 4.9 means that the increases in inputs push the production possibilities frontier outwards. Every input has a negative elasticity of the distance function and parameter estimates are significant statistically.

The largest effect caused by the livestock input (log_S) for the outputs. The second and the third important input for the outputs are the total material inputs (log_M) and the feed input (log_F). The negative trend parameter input means that there is technological improvement over the years, which push the production possibility sets outwards over the years.

The scale elasticity of the distance function, which is calculated by adding the distance elasticties with respect to the six inputs is -1.044 thus we conclude that the dairies are operation in the increasing returns to scale part of the technology on average.

The estimate of η is negative, which is suggesting significant increasing in technical efficiency over these The log-likelihood in terms of **$\gamma = \sigma_{u^2}/\sigma^2$** is 0.868 which means that 86.8% of the variation in the composite error term is due to the inefficiency component.

## 4.3 Models results comparison

The similar objective of the Data Envelopment Analysis (DEA) and the Stochastic Frontier Analysis (SFA) is that both methods are used to measure efficiency, but in a different way and under different assumptions. Like the DEA is a non-parametric method to measure the relative efficiency of a farm in the presence of multiple inputs and outputs, without knowing the production function or in other words without needing to know a functional relationship between inputs and outputs. Contrarily the SFA is a parametric method estimates a translog production function to describe the functional relationship between the multiple inputs and outputs. The translog function is quiet a flexible stochastic function because of the less data restrictions comparing the original Cobb-Douglas stochastic frontier model.

Other differences between the two methods over that one is non-parametric and doesn’t count with the statistical noise and the other is parametric and counts with that noise are the different assumptions what they use for calculating the efficiency scores. For instance in the non-parametric DEA case given the assumption that noise does not exist, it is clear that the height of the DEA frontier is biased downwards in finite samples and hence that the efficiency scores are biased upwards (Coelli et. al. 2005). The parametric SFA methods includes two random terms the symmetric error (v) and a non-negative random inefficiency variable (u). The SFA assumes that total production deviates from the optimal production by a random noise and an inefficiency component. The noise component is assumed to follow a symmetric normal distribution around the frontier that captures random disturbances, thus the noise components are independently and identically distributed normal random variables with zero means and variances. The inefficiency component can be distributed as half-normally (or for instance: truncated normally, exponentially or gamma) and represents the individual firm’s deviations from the efficient frontier. The inefficiency component has similar properties as the noise component except it has a nonzero mean, but these terms are also independently and identically distributed half-normal random variables with scale parameter (Coelli et. al. 2005)..

The next tables summarises the results of the examined countries technical efficiency using DEA and SFA and assuming national and common frontier among the three countries, which can give different estimates of technical efficiency. The difference between the results might be caused by the different assumptions the two methods and the different sample size as well.

Table 4.10 presents the technical efficiency results using DEA method. There are two cases inside the analysis, which are assuming national or common frontier. Observable that in the common frontier case all the countries technical efficiency scores under variable returns to scale are lower than the common frontier cases (VRS TE). That might have caused the bigger sample size, and the assumption that all the examined countries can operate in the same environmental conditions and can employ the same technology. We can see this effect in the table 4.13 as well, where we used SFA to determine the technical efficiency scores.

*Table 4.10: Comparing summary of the technical efficiency using DEA*

| Type of the DEA frontier | Country | CRS TE | VRS TE | Scale | Observations |
|---|---|---|---|---|---|
| **NATIONAL frontier** | **Germany** | 0.80 | **0.83** | 0.96 | 982 |
| | **The Netherlands** | 0.89 | **0.92** | 0.96 | 178 |
| | **Hungary*** | 0.86 | **0.90** | 0.96 | 23.4 |
| **COMMON frontier** | **Germany** | 0.75 | **0.77** | 0.97 | 982 |
| | **The Netherlands** | 0.88 | **0.90** | 0.98 | 178 |
| | **Hungary** | 0.76 | **0.84** | 0.92 | 27 |

Note: CRS TE = technical efficiency from DEA constant returns to scale; VRS TE = technical efficiency from DEA variable returns to scale; scale = scale efficiency = CRS TE / VRS TE
Note: * The Hungarian data’s are unbalanced assuming national frontier from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

Table 4.10 shows that using DEA and assuming the national frontier, the highest efficiency score is 0.92 for the Netherlands, the second is Hungary (0.9), but the sample size in that case is quiet low So that result is not trustable, because it might represent just the biggest farms which are using the best technology in Hungary. Assuming the common frontier makes our technical efficiency scores lower, earlier we mentioned the reasons of it.

*Table 4.11: Comparing summary of the technical efficiency using SFA*

| Type of the SFA frontier | Country | TE | Std. Dev. | # of Obs. |
|---|---|---|---|---|
| **NATIONAL frontier** | **Germany** | **0.76** | 0.12 | 982 |
| | **The Netherlands** | **0.84** | 0.10 | 178 |
| | **Hungary*** | **0.68** | 0.03 | 23.4 |
| **COMMON frontier** | **Germany** | **0.67** | 0.11 | 982 |
| | **The Netherlands** | **0.84** | 0.10 | 178 |
| | **Hungary** | **0.61** | 0.13 | 27 |

Note: TE = mean of the technical efficiency from SFA; Std. Dev.: standard deviation; # of Obs.: Number of observations
Note: * The Hungarian data's are unbalanced assuming national frontier from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

Table 4.11 presents the same technical efficiency score like the Table 4.10, but using the SFA method. In that case the highest score is in the Netherland with 0.84; the second is Germany with 0.76 and follows Hungary (0.68). Because of the small sample size for Hungary in assuming the national frontier we use data from 2001 to 2008 instead of 2001 to 2005. As we mentioned before the common frontier technical efficiency scores are lower in that case as well as using the DEA method.

Table 4.12 presents the two methods results. The DEA technical efficiency results are higher than the SFA results, because the DEA create a broken frontier line instead of a functional curve what a SFA method assumes. The different between the two kinds of efficiency scores are between 7-15%, except Hungary, where the different are huge around 30-38%.

*Table 4.12: Comparing summary of the technical efficiency using DEA and SFA*

| Type of the frontier | Country | DEA VRS TE | SFA TE | *# of Obs.* |
|---|---|---|---|---|
| **NATIONAL frontier** | Germany | 0.83 | 0.76 | *982* |
| | The Netherlands | 0.92 | 0.84 | *178* |
| | Hungary* | 0.90 | 0.68 | *23.4* |
| **COMMON frontier** | Germany | 0.77 | 0.67 | *982* |
| | The Netherlands | 0.90 | 0.84 | *178* |
| | Hungary | 0.84 | 0.61 | *27* |

Note:* The Hungarian data's are unbalanced assuming national frontier from 2001-2008
Source: Own calculation based of the FADN database 2001-2005.

Summarising of the results we can say, that the Netherlands has highest technical efficiency between 84-92%, the second is Germany with 76-83% and Hungary with 68-90%. But the Hungarian results are less trustable than the others, because of the low sample size.

If we want to eliminate the low sample size effect, we can assume a common frontier, which decrease the efficiency scores a bit, and makes the Hungarian results more reliable.

If we combine the two methods and the two kinds of frontiers results and take the averages of the two methods and the two kinds of frontier, we get that the most efficient farms are in the Netherlands with 87-88% efficient. The German farms are 72-80% efficient. The Hungarian farms are 73-79% efficient.

# 5 Discussion and conclusions

## 5.1 Discussion

The methods in this research were suitable and the most widely used methods to compare dairy farms efficiency for farm and national level. The DEA and the SFA methods that have been used in this research help to measure technical efficiency with using multiple outputs and multiple inputs. From the literature review we saw that it is hard to compare countries using just the partial productivity indexes, where we can examine the farms efficiency in just one dimension. Using DEA or SFA methods, we can examine the farm's technical efficiency in a multidimensional level.

The database of the research has been collected by the European Union's FADN system from 2001 to 2005 and from 2001 to 2008 for Hungary. The small number of observations per year is the reason why the Hungarian database continues more years in the sample. Thus the time horizon of the data is 5 or 8 years, but it can be longer like 10 or 20 years to get more valid results for the comparison. The number of dairy farms in the sample per year is 982 for Germany, 178 for the Netherlands and 23 for Hungary. In the future research it is desirable to increase the numbers of Hungarian dairy farms in the sample as high as the other countries farms number to get more clear view about their management for the comparison. But in the present FADN database for Hungary is not that wide about the specialised dairy farmers. On the other hand it is also possible that the Hungarian farms are not as specialised only for milk production as the Dutch or the German farms.

It appears from the results that the most efficient farms are in the Netherlands, and then Germany and Hungary follow. But the country results from the two methods are sometimes showing some difference. The non-parametric DEA and the parametric SFA methods measure the same technical efficiency, but under different assumptions, so the small differences between the results might be caused by the different method assumptions. In our research the difference between the two methods was 10-15 percent from Germany and The Netherlands and because of the small sample size 30-35 percent from Hungary. One important assumption arise using common frontier case that the three counties can access to the most efficient technology.  So under these assumptions our results indicate that the Dutch farmer are use their lands more intensively and efficiently than the other countries, thus the amount of lands has larger impact on outputs. The results also suggests that Hungarian diary sector needs some technological development to improve their efficiency, because of the old machinery and buildings so they need more capital to do that, thus these inputs have more effects for the outputs.

We can see in our database, that there are only few specialised big farms comparable to the Dutch and German farms, that’s one reason for the  small Hungarian sample. Although we can see that the farms are relatively efficient in the Hungarian sample comparing their national frontier. Nevertheless to get a better view about the break points of the different countries dairy efficiency, we need to make a SWOT (strength, weakness, opportunity, threats) analysis or examine allocative efficiency for their dairy sector, which require more time, capital and more experts opinions. Thus this can be a good topic for future research.

Directions for the future research can be also to estimate allocative efficiency models where the different countries, different inputs and outputs prices are also play an important role to compare efficiency among countries. Unfortunately the FADN database directly cannot contain information about prices, but indirectly we can calculate it. These analyses needs more time and more complicated model to estimate the frontiers. To get better view about the dairy sector efficiency in the future we need to analyse other important countries or sectors (feeding industry, plan cultivating sectors) which play important role of the sector or the examined country import-export market and use other methods to measure efficiencies like the total factor productivity (TFP) indexes.

The usability of these methods for other country, region sector is possible, if they have proper data for the analysis. The method is available to compare not just countries but regions inside the counties. The adaptability of this model is wide so we can analyse different sectors in the agriculture and different industrial sectors as well.

## 5.2 Conclusions

The first objective of the research is to measure dairy farms efficiency in Hungary, Germany and The Netherlands. The second objective is to compare parametric and non-parametric methods of efficiency measurement in practise.

First we compare the three countries partial efficiency indexes, which mainly comparing ratio of one input and one output. According to the results we can establish the dairy sector characteristic of the three countries. The biggest milk producer is Germany; the smallest is Hungary among the three countries. About the applied technology, the Hungarian dairy sector are land and labor extensive in contrast to the Dutch dairy sector which are land and labor intensive. This intensive farming practices can involve very large numbers of animals raised on limited land which require large amounts of feed, water and medical inputs. The German dairy sector about the land and labor are somewhere in the middle of the other two examined countries.

So far the measuring of the inputs and the outputs was carried separately, the next step was measuring the efficiency performance with respect to all inputs and all output called *"multiple inputs and output measuring".* The non-parametric DEA and the parametric SFA methods that have been used in this research help to measure technical efficiency with using multiple outputs and multiple inputs. The main difference of the two methods that the DEA measures the relative efficiency of a farm in the presence of multiple inputs and outputs, without knowing the functional relationship between inputs and outputs and given the assumption that statistical noise does not exist. Contrarily the SFA estimates a translog production function to describe the functional relationship between the multiple inputs and outputs and make allowance for the statistical noise and inefficiency effect. Because of the different model assumptions the DEA efficiency scores are biased upwards comparing to the SFA results.

We used two outputs in our models, the revenues from cow's milk production and the revenues from other outputs. For the better estimation to account for the dependence of revenues on inflation, the output revenues and the inputs are deflated with country-wide price indices for each category of products. The analysis used six deflated inputs categories, which cover the whole input side of the dairy business. These categories were the following: capital, labor, land, total material inputs, livestock and purchased feed.

The European Union's FADN database has been used for this research which contains data from 2001 to 2005 and from 2001 to 2008 for Hungary, because of the small sample size. The number of dairy farms in the sample per year was 982 for Germany, 178 for the Netherlands and 23 for Hungary. We define specialised dairy farm like those dairy farms, whose revenues from cow's milk production are at least 75% of their total revenues for every year.

It appears form the results that the Netherlands has highest technical efficiency; the second is Germany and Hungary. But the Hungarian results are less trustable than the others, because of the low sample size. Eliminating the low sample size effect with assuming a common frontier, which decrease the efficiency scores a bit, and it makes the Hungarian results more reliable. If we combine the two methods and the two kinds of frontiers results and take the averages of the two methods and the two kinds of frontier, we get that the most efficient farms are in the Netherlands after Germany and Hungary follows.

We can assume that if the quota system abolished and assuming a common price for milk in EU, only the efficient farms will survive the higher competition among the countries. In our case the Dutch farms are the most efficient, thus probably they will increase their production after the quota system. But because the size of the country we cannot expect dramatic changes in the European Dairy market. The Germans farms efficiency is lower, although their dairy sector size is bigger than the other two countries, so we won't expect high increase about the dairy supply. The Hungarian dairy sector is not as efficient as the Dutch, and the size of the sector is also small among the European countries, thus if they want to survive the quota system demolishing, they have to increase their efficiency.

# References


- Aigner, D.J., and S.F. Chu 1968, "On Estimating the Industry Production Function", American Economic Review, 58, 826-839.
- Aigner, D.J., C.A.K Lovell and P. Schmidt 1977, "Formulation and Estimation of Stochastic Frontier Production Function Models", Journal of Econometrics, 6, 21-37.
- and productivity analysis Second Edition, Springer Science-i-Business Media, Inc.USA,
- Banker, R.D., A. Chames and W.W. Cooper (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", Management Science, 30, 1078- 1092.
- Battese, G. E. and Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis, 3(1-2):153–169.
- Bauer, P.W., Berger, A.N., Ferrier, G.D., and Humphrey, D.B. (1998). Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency models. Journal of Economics and Business 50: 85–114.
- Begum I.A., Buysse J., Alam M. J.,Huylenbroeck G. V: 2009.; An application of Data Envelopment Analysis (DEA) to Evaluate Economic Efficiency of Poultry Farms in Bangladesh; presentation at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009
- Bravo-Ureta, B. E., V. H. Moreira, A. A. Arzubi, E. D. Schilder, J. Alvarez, and C. Molina. 2008. Technological change and technical efficiency for dairy farms in three countries of South America. Chilean Journal of Agricultural Research 68:360–367.
- Cabrera V. E., D. Solís, and J. del Corral 2010. “Determinants of technical efficiency among dairy farms in Wisconsin” Journal of Dairy Science 93:387–393
- Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2:429–444.
- Coelli T., Rao D.S.P., O’Donell C.J., Battese G.E. (2005) An introduction to efficiency
- Coelli, T.J., and S. Perelman (1996), "Efficiency Measurement, Multiple-output Technologies and Distance Functions: With Application to European Railways", *CREPP Discussion Paper* no. 96/05, University of Liege, Liege.
- Emvalomatis, G. 2010: Distance Functions, Inside document for the Business Economics Group, Wageningen University, Wageningen, The Netherlands
- EUROSTAT 2010. http://epp.eurostat.ec.europa.eu/portal/page/portal/agriculture/data/database
- EUROSTAT 2010a. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-08-115/EN/KS-SF-08-115-EN.PDF
  EUROSTAT 2010b. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-09-007/EN/KS-SF-09-007-EN.PDF
- EUROSTAT 2010c. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-09-039/EN/KS-SF-09-039-EN.PDF
- FADN 2010a.: http://ec.europa.eu/agriculture/rica/index.cfm

- FADN 2010b.: http://ec.europa.eu/agriculture/rica/concept_en.cfm
- FADN 2010c..:http://ec.europa.eu/agriculture/rica/methodology3_en.cfm
- FADN REPORT 2010 http://ec.europa.eu/agriculture/analysis/fadn/reports/sa0207_milk.xls
- FAOSTAT 2010. http://faostat.fao.org/site/569/DesktopDefault.aspx?PageID=569#ancor
- Farrell, M. J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A 120(3):253–281.
- Heshmati, A., and S. Kumbhakar. 1994. Farm heterogeneity and technical efficiency: Some results from Swedish dairy farms. Journal of Productivity Analysis 5:45–61.
- Hormazábal and Wyngard 2007. Frontier methodologies for the determination of efficiencies in distribution costs, Ingeniare. Fondecyt 1040339 and Universidad de Tarapacá Project 8721-05 and 8723-06, Revista chilena de ingeniería, vol. 15 Nº 3, 2007, pp. 220-226
- Jaforullah, M., and J. Whiteman. 1999. Scale efficiency in the New Zealand dairy industry: A non-parametric approach. Aust., Journal of Agricultural and Resource Economics 43:523–541.
- Kasman, A. and Turgutlu, E 2007.: A Comparison of Chance-constrained DEA and Stochastic Frontier Analysis: An Application to the Turkish Life Insurance Industry, 8. Türkiye Ekonometri ve İstatistik Kongresi 24-25 Mayıs 2007 – İnönü Üniversitesi Malatya http://web.inonu.edu.tr/~eisemp8/bildiri-pdf/kasman-turgutlu.pdf
- Kumbhakar, S., and K. Lovell. 2000. Stochastic Frontier Analysis. Cambridge Univ. Press, Cambridge, UK.
- Lawson, L. G., J. Bruun, T. Coelli, J. F. Agger, and M. Lund. 2004. Relationships of efficiency to reproductive disorders in Danish milk production: A stochastic frontier analysis. Journal of Dairy Science 87:212–224.
- Meeusen, W. and J. van den Broeck 1977, "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error", International Economic Review, 18, 435-444.
- Murova, O., and B. Chidmi. 2009. Impacts of Federal Government programs and specific farm variables on technical efficiency of dairy farms. Paper presented at the 2009 Southern Agricultural Economics Association Annual Meeting, Atlanta, Georgia (available at http://purl.umn.edu/46822).
- Solís, D., B. Bravo-Ureta, and R. Quiroga. 2009. Technical efficiency among peasant farmers participating in natural resource management programs in Central America. Journal of Agricultural Economics 60:202–219.
- Stokes, J. R., P. R. Tozer, and J. Hyde. 2007. Identifying efficient dairy producers using data envelopment analysis. Journal of Dairy Science 90:2555–2562.
- Tauer, L. W. 1998. Productivity of New York dairy farms measured by nonparametric Malquist indices. Journal of Agricultural Economics 49:234–249.
- Tauer, L. W., and A. K. Mishra. 2006. Dairy farm cost efficiency. Journal of Dairy Science 89:4937–4943.

# APPENDIX

*A1.: Summary of the input slacks in Germany*

| Year/input | Capital | Labor | Land | Material inputs | Livestock | Feed |
|---|---|---|---|---|---|---|
| 2001 | 31252.17 | 436.95 | 9.06 | 758.48 | 2.33 | 443.86 |
| 2002 | 28104.32 | 322.46 | 9.54 | 1030.04 | 3.37 | 586.22 |
| 2003 | 28145.15 | 285.91 | 11.07 | 743.00 | 3.16 | 588.50 |
| 2004 | 24875.65 | 322.56 | 9.75 | 1041.25 | 2.64 | 444.02 |
| 2005 | 26027.69 | 248.14 | 10.04 | 1175.10 | 3.67 | 331.25 |
| Average | 27681.00 | 323.20 | 9.89 | 949.57 | 3.03 | 478.77 |

Source: Own calculation based of the FADN database 2001-2005.

*A2..: Summary of the input slacks in The Netherlands*

| Year/input | Capital | Labor | Land | Material inputs | Livestock | Feed |
|---|---|---|---|---|---|---|
| 2001 | 25557.39 | 585.04 | 4.19 | 2712.19 | 4.23 | 1172.14 |
| 2002 | 12122.71 | 343.76 | 3.19 | 2915.12 | 2.67 | 709.959 |
| 2003 | 20649.32 | 407.53 | 2.30 | 1821.49 | 2.25 | 1341.44 |
| 2004 | 21983.88 | 290.18 | 2.52 | 2189.72 | 2.18 | 885.51 |
| 2005 | 16965.95 | 219.15 | 3.09 | 2326.22 | 2.32 | 665.808 |
| Average | 19455.85 | 369.13 | 3.06 | 2392.95 | 2.73 | 954.97 |

Source: Own calculation based of the FADN database 2001-2005.

*A3.: Summary of the input slacks in Hungary*

| Year/input | Capital | Labor | Land | Material inputs | Livestock | Feed |
|---|---|---|---|---|---|---|
| 2001-2004 | 11026.53 | 1677.07 | 22.75 | 592.35 | 5.01 | 562.087 |
| 2005-2008 | 15346.60 | 841.44 | 8.88 | 1527.74 | 5.85 | 2005.45 |
| Average | 13186.57 | 1259.25 | 15.82 | 1060.04 | 5.43 | 1283.77 |

Source: Own calculation based of the FADN database 2001-2008.

*A 4.: Estimates of the Time-varying SFA model's parameters*

| | Germany | | | The Netherlands | | | Hungary* | | |
|---|---|---|---|---|---|---|---|---|---|
| **log_cmilk** | **Coef.** | **Std. Err.** | **p-value** | **Coef.** | **Std. Err.** | **p-value** | **Coef.** | **Std. Err.** | **p-value** |
| log_oth | 0.189 | 0.003 | 0.000 | 0.099 | 0.010 | 0.000 | 0.327 | 0.021 | 0.000 |
| log_K | -0.054 | 0.006 | 0.000 | -0.043 | 0.010 | 0.000 | -0.139 | 0.033 | 0.000 |
| log_L | -0.060 | 0.011 | 0.000 | -0.022 | 0.017 | 0.195 | 0.083 | 0.060 | 0.167 |
| log_A | -0.047 | 0.013 | 0.000 | -0.158 | 0.025 | 0.000 | -0.115 | 0.050 | 0.021 |
| log_M | -0.210 | 0.011 | 0.000 | -0.092 | 0.021 | 0.000 | -0.228 | 0.066 | 0.001 |
| log_S | -0.445 | 0.015 | 0.000 | -0.520 | 0.030 | 0.000 | -0.527 | 0.080 | 0.000 |
| log_F | -0.156 | 0.006 | 0.000 | -0.193 | 0.015 | 0.000 | -0.122 | 0.030 | 0.000 |
| log_KK | -0.008 | 0.003 | 0.011 | 0.039 | 0.013 | 0.002 | -0.022 | 0.024 | 0.362 |
| log_KL | -0.022 | 0.014 | 0.133 | 0.030 | 0.026 | 0.247 | -0.020 | 0.074 | 0.789 |
| log_KA | -0.021 | 0.015 | 0.177 | 0.000 | 0.038 | 0.991 | -0.047 | 0.063 | 0.458 |
| log_KM | 0.042 | 0.012 | 0.000 | -0.062 | 0.032 | 0.052 | 0.193 | 0.072 | 0.007 |
| log_KS | 0.014 | 0.017 | 0.386 | -0.069 | 0.047 | 0.141 | -0.071 | 0.087 | 0.415 |
| log_KF | -0.001 | 0.005 | 0.909 | 0.030 | 0.023 | 0.195 | -0.048 | 0.032 | 0.134 |
| log_LL | -0.017 | 0.019 | 0.352 | -0.017 | 0.019 | 0.370 | 0.069 | 0.078 | 0.374 |
| log_LA | -0.013 | 0.029 | 0.655 | -0.060 | 0.066 | 0.367 | 0.143 | 0.094 | 0.129 |
| log_LM | 0.076 | 0.028 | 0.006 | -0.172 | 0.057 | 0.003 | -0.101 | 0.129 | 0.436 |
| log_LS | -0.017 | 0.038 | 0.649 | 0.213 | 0.080 | 0.007 | -0.378 | 0.174 | 0.030 |
| log_LF | 0.028 | 0.015 | 0.056 | 0.058 | 0.049 | 0.234 | 0.187 | 0.063 | 0.003 |
| log_AA | 0.055 | 0.021 | 0.009 | 0.071 | 0.057 | 0.216 | -0.005 | 0.043 | 0.908 |
| log_AM | -0.103 | 0.028 | 0.000 | 0.070 | 0.077 | 0.363 | 0.125 | 0.104 | 0.227 |
| log_AS | -0.004 | 0.041 | 0.916 | -0.411 | 0.111 | 0.000 | -0.348 | 0.153 | 0.023 |
| log_AF | 0.032 | 0.013 | 0.018 | 0.215 | 0.061 | 0.000 | 0.113 | 0.071 | 0.114 |
| log_MM | 0.082 | 0.016 | 0.000 | 0.081 | 0.043 | 0.059 | -0.055 | 0.094 | 0.562 |
| log_MS | -0.270 | 0.036 | 0.000 | 0.025 | 0.097 | 0.797 | 0.174 | 0.213 | 0.415 |
| log_MF | 0.040 | 0.012 | 0.001 | 0.005 | 0.051 | 0.914 | -0.205 | 0.075 | 0.006 |
| log_SS | 0.163 | 0.034 | 0.000 | 0.283 | 0.090 | 0.002 | 0.216 | 0.136 | 0.113 |
| log_SF | -0.038 | 0.016 | 0.017 | -0.285 | 0.086 | 0.001 | 0.097 | 0.084 | 0.245 |
| log_FF | -0.027 | 0.002 | 0.000 | 0.036 | 0.023 | 0.124 | -0.035 | 0.014 | 0.012 |
| log_otot | 0.055 | 0.000 | 0.000 | 0.034 | 0.007 | 0.000 | 0.062 | 0.003 | 0.000 |
| log_Ko | -0.008 | 0.004 | 0.032 | -0.021 | 0.013 | 0.108 | -0.023 | 0.016 | 0.161 |
| log_Lo | -0.034 | 0.006 | 0.000 | 0.013 | 0.023 | 0.567 | 0.015 | 0.020 | 0.434 |
| log_Ao | 0.025 | 0.007 | 0.000 | 0.014 | 0.028 | 0.614 | 0.000 | 0.017 | 0.986 |
| log_Mo | 0.031 | 0.008 | 0.000 | -0.008 | 0.026 | 0.748 | 0.011 | 0.022 | 0.618 |
| log_So | -0.026 | 0.010 | 0.007 | 0.016 | 0.038 | 0.670 | -0.013 | 0.031 | 0.667 |
| log_Fo | 0.002 | 0.003 | 0.589 | 0.002 | 0.020 | 0.919 | -0.014 | 0.012 | 0.231 |
| trend | -0.016 | 0.002 | 0.000 | -0.020 | 0.003 | 0.000 | -0.004 | 0.010 | 0.693 |
| trend_trend | -0.005 | 0.001 | 0.000 | 0.001 | 0.001 | 0.260 | -0.005 | 0.005 | 0.265 |
| otht | 0.000 | 0.001 | 0.526 | 0.004 | 0.003 | 0.217 | -0.001 | 0.003 | 0.800 |
| Kt | -0.003 | 0.001 | 0.074 | 0.002 | 0.004 | 0.496 | -0.002 | 0.014 | 0.879 |
| Lt | -0.009 | 0.003 | 0.007 | -0.008 | 0.006 | 0.175 | -0.010 | 0.018 | 0.572 |
| At | 0.000 | 0.003 | 0.891 | 0.001 | 0.008 | 0.910 | -0.005 | 0.020 | 0.823 |
| Mt | -0.005 | 0.004 | 0.138 | 0.000 | 0.007 | 0.973 | -0.006 | 0.024 | 0.806 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| St | 0.012 | 0.004 | 0.005 | -0.011 | 0.011 | 0.316 | 0.015 | 0.030 | 0.617 |
| Ft | 0.002 | 0.002 | 0.123 | 0.014 | 0.006 | 0.036 | -0.003 | 0.011 | 0.783 |
| constant | -0.403 | 0.012 | 0.000 | -0.421 | 0.017 | 0.000 | -0.857 | 0.639 | 0.180 |
| **$\mu$** | 0.207 | 0.033 | 0.000 | -0.070 | 0.223 | 0.753 | 0.389 | 0.635 | 0.540 |
| **$\eta$** | 0.001 | 0.006 | 0.850 | -0.044 | 0.015 | 0.003 | -0.019 | 0.077 | 0.808 |
| **$\sigma^2$** | 0.064 | 0.008 | 0.000 | 0.079 | 0.040 | 0.000 | 0.056 | 0.006 | 0.000 |
| **$\gamma$** | 0.882 | 0.015 | 0.000 | 0.953 | 0.023 | 0.000 | 0.154 | 0.204 | 0.000 |
| **$\sigma_{u2}$** | 0.057 | 0.008 | 0.000 | 0.076 | 0.040 | 0.000 | 0.009 | 0.012 | 0.000 |
| **$\sigma_{v2}$** | 0.008 | 0.000 | 0.000 | 0.004 | 0.000 | 0.000 | 0.047 | 0.011 | 0.000 |

Note: * The Hungarian data are unbalanced from 2001-2008

Source: Own calculation based of the FADN database 2001-2005.

*A 5.: Estimates of the Time-varying SFA model's parameters assuming common frontier*

| **log_cmilk** | **Coef.** | **Std. Err.** | **p-value** |
|---|---|---|---|
| log_oth | 0.178 | 0.003 | 0.000 |
| log_K | -0.059 | 0.005 | 0.000 |
| log_L | -0.053 | 0.010 | 0.000 |
| log_A | -0.023 | 0.012 | 0.058 |
| log_M | -0.221 | 0.010 | 0.000 |
| log_S | -0.523 | 0.014 | 0.000 |
| log_F | -0.165 | 0.006 | 0.000 |
| log_KK | -0.008 | 0.003 | 0.010 |
| log_KL | -0.034 | 0.011 | 0.003 |
| log_KA | 0.006 | 0.012 | 0.606 |
| log_KM | 0.045 | 0.010 | 0.000 |
| log_KS | -0.021 | 0.015 | 0.153 |
| log_KF | 0.002 | 0.005 | 0.722 |
| log_LL | -0.008 | 0.012 | 0.518 |
| log_LA | 0.052 | 0.021 | 0.011 |
| log_LM | -0.002 | 0.021 | 0.932 |
| log_LS | -0.074 | 0.030 | 0.014 |
| log_LF | 0.055 | 0.012 | 0.000 |
| log_AA | 0.061 | 0.017 | 0.000 |
| log_AM | -0.043 | 0.023 | 0.062 |
| log_AS | -0.130 | 0.033 | 0.000 |
| log_AF | 0.036 | 0.012 | 0.003 |
| log_MM | 0.071 | 0.014 | 0.000 |
| log_MS | -0.169 | 0.033 | 0.000 |
| log_MF | 0.012 | 0.011 | 0.280 |
| log_SS | 0.230 | 0.028 | 0.000 |
| log_SF | -0.016 | 0.015 | 0.282 |
| log_FF | -0.029 | 0.002 | 0.000 |
| log_otot | 0.056 | 0.000 | 0.000 |
| log_Ko | -0.024 | 0.003 | 0.000 |
| log_Lo | -0.024 | 0.006 | 0.000 |
| log_Ao | 0.021 | 0.006 | 0.000 |
| log_Mo | -0.027 | 0.006 | 0.000 |
| log_So | 0.029 | 0.009 | 0.001 |
| log_Fo | -0.003 | 0.003 | 0.233 |
| trend | -0.019 | 0.002 | 0.000 |
| trend_trend | -0.004 | 0.001 | 0.000 |
| otht | 0.001 | 0.001 | 0.023 |
| Kt | -0.002 | 0.001 | 0.108 |
| Lt | -0.010 | 0.003 | 0.001 |
| At | -0.001 | 0.003 | 0.847 |
| Mt | -0.005 | 0.003 | 0.117 |
| St | 0.009 | 0.004 | 0.020 |

| | | | |
|---|---|---|---|
| Ft | 0.004 | 0.002 | 0.005 |
| constant | -0.527 | 0.018 | 0.000 |
| **μ** | 0.365 | 0.028 | 0.000 |
| **η** | -0.007 | 0.005 | 0.153 |
| $\boldsymbol{\sigma}^2$ | 0.062 | 0.005 | |
| **γ** | 0.868 | 0.012 | |
| $\boldsymbol{\sigma}_{u2}$ | 0.054 | 0.005 | |
| $\boldsymbol{\sigma}_{v2}$ | 0.008 | 0.000 | |

Note: * The Hungarian data are unbalanced from 2001-2008

Source: Own calculation based of the FADN database 2001-2005.