Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

**Department of Economics**

# Analyzing the technical performance of Swedish farms

- Involved in horse related activities

*Per Andersson*

Master's thesis · 30 hec · Advanced level
Agricultural Economics and Management – Master's Programme
Degree thesis No 898 · ISSN 1401-4084
Uppsala 2014

**Analyzing the technical performance of Swedish farms**

*Per Andersson*

**Supervisor:** Yves Surry, Swedish University of Agricultural Sciences, Department of Economics

**Assistant supervisor:** Hans Andersson, Swedish University of Agricultural Sciences, Department of Economics

**Examiner:** Ing-Marie Gren, Swedish University of Agricultural Sciences, Department of Economics

**Credits:** 30 hec
**Level:** A2E
**Course title:** Independent Project/Degree Project in Economics
**Course code:** EX0537
**Programme/Education:** Agricultural Economics and Management, Master's Programme
**Faculty:** Faculty of Natural Resources and Agricultural Sciences

**Place of publication:** Uppsala
**Year of publication:** 2014
**Name of Series:** Degree project/SLU, Department of Economics
**No:** 898
**ISSN** 1401-4084
**Online publication:** http://stud.epsilon.slu.se

**Key words:** *Agricultural production, economics, horse activities, parameters of technical inefficiency, stochastic frontier analysis, technical efficiency.*

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

**Department of Economics**

# Abstract

This study analyses the technical performance of Swedish farms involved in horse related activities. The technical performance is analyzed by the technical efficiency of the farms and the factors that contribute to technical efficiency. The method used in this study is that of stochastic frontier analysis, where the stochastic frontier production function is estimated in order to obtain the technical efficiency scores of the Swedish farms. The factors that contribute to technical inefficiency are then estimated simultaneously with the stochastic frontier production function in order to see the effect the factors have on the technical performance of the Swedish farms.

The relevant data has been collected from a survey implemented and supervised by LRF Konsult and designed by the department of Economics at the Swedish University of Agricultural Sciences. The complete sample data consists of 428 observations, but only 58 are used in the analysis of the technical performance. The output of the Swedish farms consist of income from agricultural production and income from horse related activities. The used by the farms are land, labor, capital and intermediate consumption. The results show that there is no farm that is neither fully efficient nor fully inefficient and where the mean technical efficiency is 0.69. Factors that do contribute to an efficient farm operation that is involved in horse related activities are farmers that receive grants. The factor age does also contribute to technical efficiency where younger farmers tend to be more efficient than older farmers do.

# Table of content

# List of figures

# List of tables

# 1 Introduction

## 1.1 Background

The use of horses has been changing throughout history and they have been of great importance for humankind (Johansson *et al.*, 2004). In the prehistoric age, the horse was mainly a source for food (ibid.) but around 4000 BC the horse was tamed and was used for pulling goods (Furugren, 1990). The first known trace of tame horses in Sweden dates back to 2700 BC (Myrdal, 2001). The early horses in Sweden were small but they were despite their size powerful and perseverant. During the Iron Age, the small Swedish horses were crossbreed with larger horses from southern Europe (Furugren, 1990).

For a long period, horses were used in warfare for pulling goods and soldiers as well as mounts (Furugren, 1990). Findings show that they were used in warfare as early as 2000 BC. As the breeding for larger and more powerful horses and the development of better accessories progressed the usage of horses shifted to forestry and farming (Johansson *et al.*, 2004). In Sweden, horses were mainly used in the northern parts and oxen were used in the middle- and southern parts (SLU, 2001). In 1850, the mechanization of agriculture developed and the demand for horses in the production was growing (Furugren, 1990). Sweden started to import more powerful and perseverant horses as the demand for such attributes increased.

The demand for horses declined as the mechanization and technical innovation in agricultural production progressed (Johansson *et al.*, 2004). During the last century, horses have gone from being a part of the production of agricultural goods to become mainly consumers of it. A sustainable population of horses needs the production of feeds, areas for grazing, equipment, and the purchase of services. Even though horses are no longer a part of the production of agricultural goods, they are producers of different services e.g. recreation, gambling and leisure-time activity.

In 2004, the Swedish Board of Agriculture did a review of the amount of horses in Sweden (from here on SJV). The review showed that there were roughly 283 100 horses in Sweden at the time (Jordbruksverket, 2011). In 2010 the amount of horses were roughly 362 700. SJV estimate that the real growth of horses in Sweden was 10-20 % between the years of 2004 and 2010. The average amount of horses per 1000 inhabitants is 39 (Jordbruksverket, 2011) which makes Sweden the country that has the most horses per 1000 inhabitants in Europe (Liljenstolpe, 2009). As for the distribution of horses in Sweden, the largest amount of horses per region is in the middle and southern regions of Sweden or close to larger cities e.g. Stockholm, Göteborg or Malmö (Jordbruksverket, 2011).

Johansson *et al*. (2004) made a study named *"The economic importance of the horse sector in Sweden"* where an input-output model was constructed which explains the direct and indirect effects the horse sector in Sweden has on other sectors and on the economy as a whole. They found that the horse sector has a direct effect on the social economy in Sweden and had an effect on the Swedish GDP. The direct effect consists of the circulation of money in the economy, tax income and employment that all originate from the horse activities created by the horse owners. The tax income is one of the largest direct effects that originate from betting on horses. The gambling is also the largest industry (in monetary value) in the horse sector (Johansson *et al.*, 2004). The effects come from the number of people that gamble, which in turn spurs the trainers, the breeding and other activities that are somehow connected to gambling. Other contributory factors that affect the economic activity in the sector are the domestically owned horses (which are the greatest in numbers), riding schools, tourism and other horse-related establishments. Despite their utilization, the horse also has a social economic effect due to its demand for forage, services, equipment and boarding.

The results from the study made by Johansson *et al.* (2004) show that 9 466 full time workers are directly dependent on the horse sector in Sweden. The largest part is dependent on the gambling industry (2 664) while the second largest is agriculture (2 180). As a large number of horse-related establishments in Sweden are located on the countryside, they create work opportunities there that in turn contribute to rural development.

The horse sector was in 2007 the fifth largest source for income in the agricultural sector in Sweden. Calculations show that the horse industry contribute with 1 500 million SEK per year to the agricultural sector, which correspond to 4 % of the total income of the agricultural sector (Tell, 2001). Out of the 1 500 million SEK, 85 % goes to farmers that produce and sell forage to the horse sector. The production of forage to the horse sector has developed over the years: from being the excess supply to the dairy production to the production of quality forage for horses only (Hedberg, 2002). More and more farmers are today producing high quality forage for horse owners (mainly trainers) who have a demand for it.

Today in Sweden there are more firms that has horses than has dairy cows (Internet, SJV, 1-2, 2013). The number of dairy cows are decreasing while the number of horses in Sweden are increasing and in 2010, the number of horses outnumbered the number of cows (Internet, SCB, 2013). A large part of the horses in Sweden, 117 000 of the total amount of horses (362 700), are found on farms (Jordbruksverket, 2011). The horses that are found on farms are not always used in the agricultural production but are owned by the farmers for recreational purposes or are boarded (Hedberg, 2002). Today many farmers in Sweden do not only devote themselves to the traditional agricultural production. Instead, they might sell forage to the horse industry, devote themselves to tourism on the countryside or provide the ability to board horses.

## 1.2 Problem

For a long time, horses have been of great importance to humankind and the society, but there have been fluctuations over time (Hedberg, 2002). The use of the horse has gone from being an important part in warfare and an essential part in the agricultural production. Today, horses are a source for recreation or are used in sports and gambling. As the horse sector has become what it is today, different studies show that there is an economic importance of the horse sector. The economic activities that the horse industry creates have effects on other sectors and industries in Sweden as well as work opportunities. The effects may differ between regions but may be the greatest on the countryside (ibid.).

Nowadays, several farmers in Sweden do not only have their agricultural production as their main source of income, but instead they combine their agricultural production with other production processes. As the demand for quality feed and concentrate feed is great in the horse sector, some farmers have chosen to combine their agricultural production with having a function of horse related activities on the farm. There are also horse owners that have started an agricultural production in order to supply their demand for feeds.

A large amount of research about horses and the horse sector are made every year (see e.g. HNS). The research is mainly done in the fields of physiology and diseases. However, foundations as well as the horse industry have lately been interested in the research of the economic perspective of the Swedish horse sector. Studies have been made regarding the socioeconomic importance of the horse sector in Sweden using an input-output model (see Johansson *et al.*, 2004).

It is evident that the horse sector is of great economic importance for the agricultural sector in Sweden and where the demand for feed, boarding and grazing land contribute the most to the economy of the agricultural sector. Due to the demand from the horse sector, many Swedish farmers that have engaged themselves in a multifunctional[1] agricultural production. The multifunctional production might have a positive economic effect on the farmers and in some cases; it could have a positive effect for the areas where the farm is located (due to externalities of the goods produced/provided). However, there might be circumstances where the multifunctional production may not be efficient due to inefficient use of its inputs for the production processes.

## 1.3 Aim

The aim of this study is to analyze the performance of Swedish farms that are engaged in horse related activities. The performance of the Swedish farms will be analyzed with the estimated technical efficiency score of each farm. Further, the factors that contribute to the technical efficiency of the farm will be estimated and presented. The questions that are addressed in this study are as follows:

- What is the technical efficiency of farms that have outputs originating from agricultural production and from horse activities?
- What are the factors that contribute to an efficient farm operation that is involved in both agricultural production and horse related activities?

The original aim of this study was to analyze the performance of Swedish farms that have a horse boarding operation beside their agricultural production. However, this was later changed as it was hard to capture the income and costs that specifically originated from horse boarding as many of the farms are engaged in several horse related activities.

## 1.4 Delimitations

There are two different ways of measuring technical efficiency. The first is the data envelopment analysis (DEA) and the second is the stochastic frontier analysis (SFA). The difference between them is that DEA does not use a functional form[2] and is deterministic (Bravo-Ureta *et al.*, 2007), while SFA needs, a functional form specified in the model but is not deterministic[3].

[1] "*Multifunctionality refers to the fact that an economic activity may have multiple outputs and, by virtue of this, may contribute to several societal objectives at once. Multifunctionality is thus an activity oriented concept that refers to specific properties of the production process and its multiple outputs.*" (OECD, 2001, p.11)
[2] A functional form could be either Cobb-Douglas, translogarithmic or other functional form.
[3] The difference between deterministic and stochastic is that the stochastic model allows for statistical "noise" in the data that is not captured in the deterministic model. This will be further explained in chapter 2.

In order to answer the aim, the SFA method will be used. Under the SFA, several models could be used. In this study, the use of the production frontier model will be used because of multiple inputs and multiple outputs. Multiple inputs are capital, labor, land and intermediate consumption while multiple outputs correspond to output from agricultural production and output from horse related activities.

Several statistical programs could be used in order to calculate the technical efficiency. This thesis will use the program called time series processor (TSP). Although the name of the program, it is also used for cross-sectional data and panel data (Internet, TSP, 2013). The commands for the econometric modeling are written in code, either Notepad or Microsoft Word. The program has more than 2000 installations for econometric estimation (Internet, TSP, 2013). Even though the econometric modeling is written in code, there are complete models that have been made previously uploaded on TSP's homepage.

This thesis will follow the example of a production frontier model designed in TSP's user guide, which has been applied for Battese and Coelli's paper: "*Prediction of firm-level technical efficiencies with generalized frontier production function and panel data*". The model has been modified in such a way that it does not consider panel data, but cross-sectional data. The factors of technical inefficiency is estimated with the help of FRONTIER 4.1.

## 1.5 Outline

The outline of the master thesis is presented in figure 1 below. The method used in this thesis will be presented in chapter two. There will be a short introduction and history of SFA, and an explanation of the fundamentals behind technical efficiency, along with graphical and mathematical presentations. There will also be a subchapter presenting previous studies made with the method of SFA based on a meta-regression analysis done by Bravo-Ureta *et al.* (2007).
In chapter three the data used in this study will be discussed. First, the survey that has been the source for the data will be presented and how much data has been collected from the survey. Secondly, there will be an analysis of the data of how many observations that will be used in the estimation of the technical efficiency as well as statistics of the variables. Thirdly, there will be a comparison between the data used in this study and a similar survey made by the Swedish board of Agriculture. The comparison is made in order to if there are any large deviations in the data set used in this study or there is consistency between the two data sets.

In chapter four, the results from TSP will be presented and analyzed. The parts that will be analyzed are the Cobb-Douglas production function, the translogarithmic production function and their technical efficiency scores. There will also be an analysis of the factors that either contributes or not to the technical inefficiency of the farm. The estimation of the factors is essential in order to answer the aim of this thesis.

In chapter five there will be a brief discussion regarding the data used in this thesis and the results obtained. Chapter six will contain the answer to the questions asked that are related to the purpose of this thesis as well as concluding remarks regarding the study made.

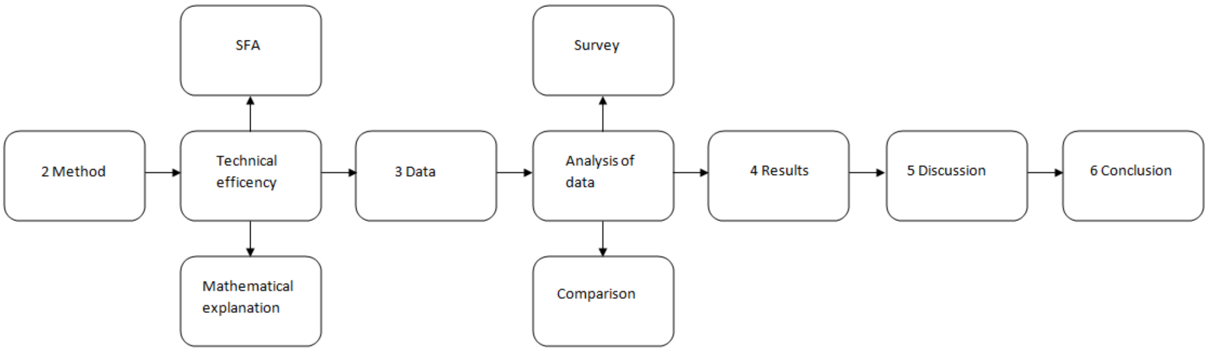



*Figure 1. Outline of thesis. (Own adaptation).*

# 2 Method

In the following chapter, the theoretical perspectives of the method will be presented. There are many articles explaining the method of stochastic frontier analysis as well as applied work using this method. The main source of inspiration and foundation for the chapter will be the book *Stochastic frontier analysis* by Kumbhakar and Lovell (2000).

## 2.1 Stochastic frontier analysis

The stochastic frontier approach was founded almost simultaneously by two groups of researchers in two separate papers by Aigner *et al.* (1977), and Meeusen and van den Broeck (1977). Stochastic frontier analysis is a method used in order to measure the efficiency of a firm or organization that uses one or multiple inputs to produce one or multiple outputs (Kumbhakar and Lovell, 2000).

There are different models that can be used to measure efficiency and which serve different purposes for the user (Kumbhakar and Lovell, 2000). There are cost-, profit- and technical efficiency. Beside technical efficiency, the user is able to measure the cost- and profit inefficiency of a firm or an organization with the use of SFA.

### 2.1.1 Technical efficiency

In 1957, Farrell introduced the econometric modeling of technical efficiency measured by the use of production functions (or frontiers). In order to explain technical efficiency, Farrell assumed for simplicity that a firm use two inputs for the production of one output and that the condition of constant returns to scale holds. He further assumed that the input-per-unit-of-output would be above the isoquant.

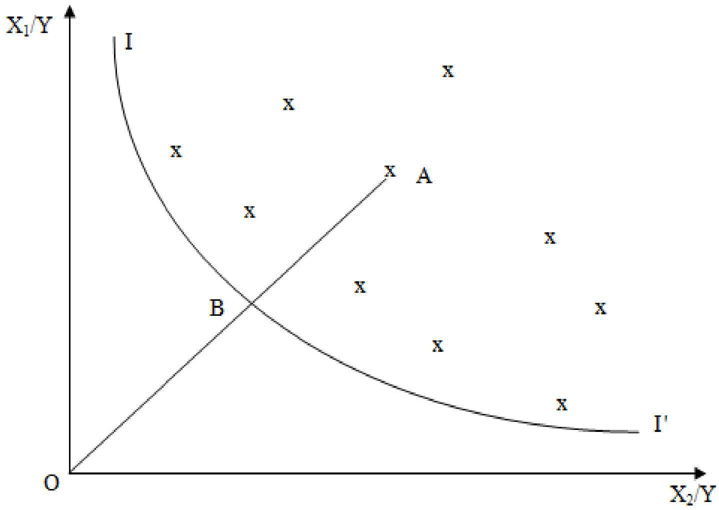



*Figure 2. Technical efficiency of firms in relative input space (Battese, 1992:186).*

In figure 2 the isoquant is represented as $II'$. The $x$'s represents the observed input-per-unit-of-output of all the observed firms, where point $A$ represents an input-per-unit-of-output for a given firm. The isoquant correspond to the efficient usage of the two inputs used to produce the output $Y$ (Farrell, 1957). Point $B$ in the diagram represents a firm that has the same input-per-unit-of-output ratio as the firm at point *A*, but which is efficient as it produces the same amount of output as point *A* but with less inputs. Farrell observed that the ratio of *OB/OA* to be the technical efficiency of the firm that has an input-per-unit-of-output ratio corresponding to point *A*.

The work of Farrell (1957) has evolved and a presentation of the production frontier more commonly used is presented in figure 3 below. On the X-axis, the vector of inputs *X* are represented which is used in order to produce output *Y* on the Y-axis (Battese, 1992). The difference between the diagram below and the previous diagram is that the input-output values, the *x*'s, are beneath the production frontier instead of above the isoquant. When the input-output values are below the production frontier, the firms do not produce the maximum amount of outputs by the given inputs (Battese, 1992). If they would produce the maximum capable output, their input-output value would be on the production frontier. The reason for the input-output values not being on the frontier may be the shortage of the available technology for the firms.

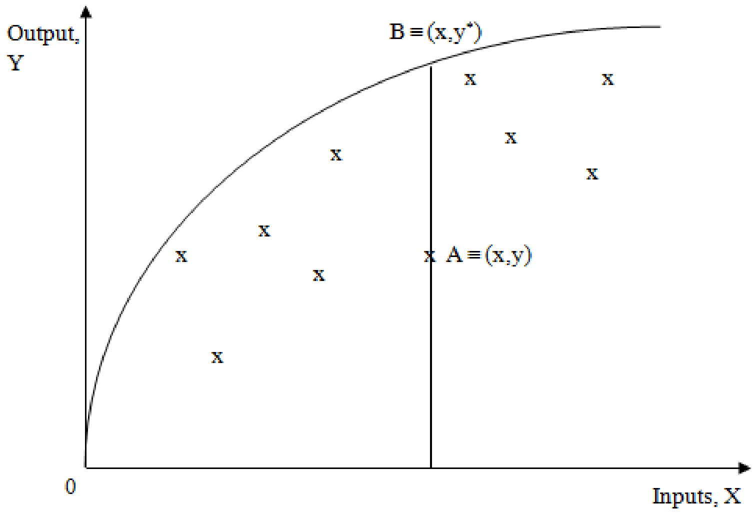



*Figure 3. Technical efficiency of firms in input-output space (Battese, 1992:187).*

The technical efficiency of a firm at point *A* equals $y/y^*$ (Battese, 1992). At point *A*, a firm produces output *y* with inputs *x*. Point *B* equals the frontier output $y^*$ which is associated with the level of inputs of the firm. Thus, technical efficiency of a firm equals the ratio between the actual output given its inputs and the frontier output, being the output that the firm could produce if the right technology were available.

The production frontier or function that is presented in figure 3 is the *deterministic* production function. The deterministic production function is not often used when it is subject to empirical analysis or when firms produce two or more outputs (Kumbhakar and Lovell, 2000). The model could be used when faced with several outputs, but the outputs have to be homogenous so that they may be aggregated into a single output. Instead of using the deterministic production function, a researcher can use the *stochastic* production frontier. The difference between them is that they observe the technical efficiency differently (Bravo-Ureta *et al.*., 2007). The deterministic assumes that the inefficiency relates to the deviation from the production frontier while the stochastic approach let statistical noise explain part of the deviation as well a part of the technical efficiency.

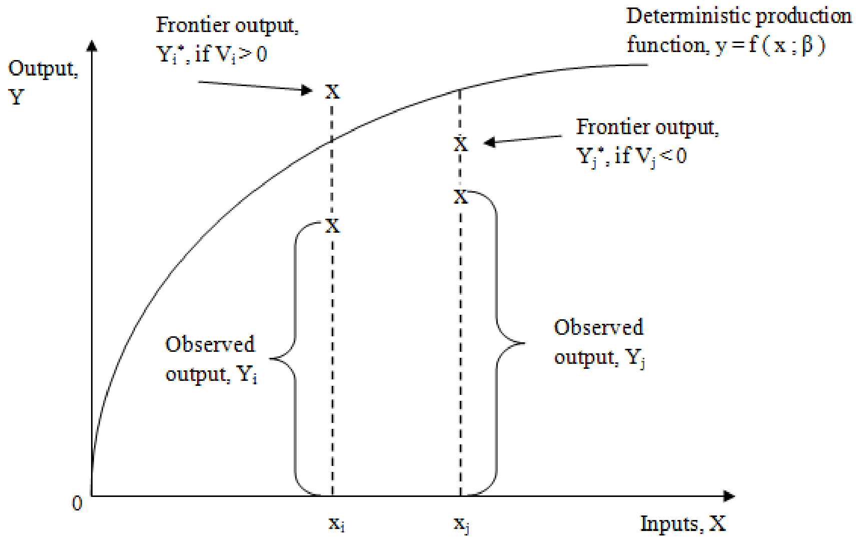



*Figure 4. Stochastic frontier production function (Battese, 1992:191).*

The stochastic frontier approach allows for statistical noise, which implies that favorable or unfavorable conditions for the firms are captured. This is showcased in figure 4 above. Firm *i* has a frontier output that is above the deterministic production function due to the error component $v_i$ being larger than zero (implying favorable conditions) (Battese, 1992). For firm *j* it is the opposite, where it has a frontier output less than the deterministic production function due to the error component being less than zero.

As is viewed in figure 4, the observed output for the two firms *i* and *j* are less than that of their frontier output. However, their unobservable frontier production outputs would lie close to the individual firms' deterministic production function (Battese, 1992). The calculation of the technical efficiency for the deterministic and stochastic frontier models are the same, but their values will be different.

Figure 4 show that the distance between the frontier output and the observed output is different for firms *i* and *j*. For firm *j*, the distance to the frontier output is less than to the deterministic production function, this implies that the firm would be considered more efficient in the stochastic frontier model than in the deterministic production frontier (Battese, 1992). The opposite holds for firm *i*, where the distance to the frontier output is larger than the distance to the deterministic production function. The value of the technical efficiencies for the firms will be less when using a deterministic frontier than a stochastic frontier. Because of a different estimation, no values will ever exceed the deterministic frontier and thus the estimation of the technical efficiencies will be less.

### 2.1.2 Stochastic output distance function

Kumbhakar and Lovell (2000) write in their book that either the input distance function or the output distance function is the best alternative to assess the technical efficiency when there are multiple outputs. Either the input or the output distance function is closely related to the production technology of a firm and is therefore used as a tool for assessing the technical efficiency. Despite this, it is rarely used to econometrically assess the technical efficiency of a firm that uses multiple inputs to produce multiple outputs (Kumbhakar and Lovell, 2000).
Two problems could arise when trying to measure the technical efficiency using stochastic distance function model: 1) the recognition of the dependent variable is tough when there are

multiple outputs and 2) the regressors might not act endogenously (Kumbhakar and Lovell, 2000). However, these problems can be resolved and the use of the stochastic distance function model will work similarly as a production frontier model, but with multiple outputs.
An output distance function looks closer at the measurement of how the distance from the producer and its inputs are from the production possibilities the producer has (Kumbhakar and Lovell, 2000). The output distance function explains how a given input vector can produce a minimum amount of an output vector before it is deflated.

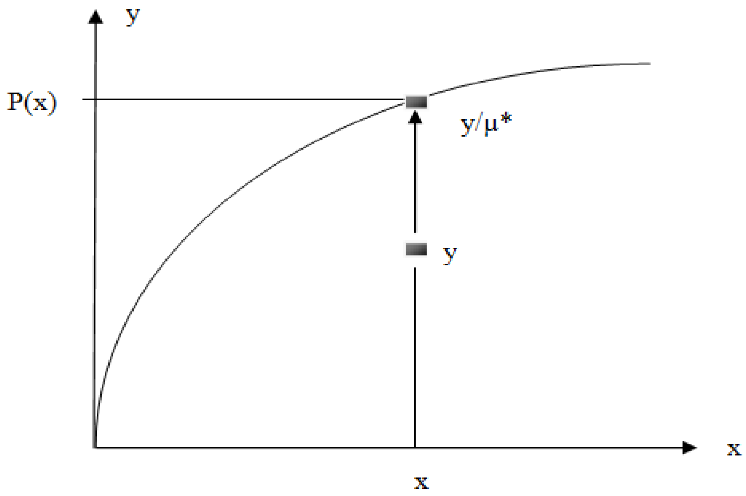



*Figure 5. An output distance function (Kumbhakar and Lovell, 2000:31).*

Figure 5 present an output distance function where *y* equals output and *x* equals input. The curved line corresponds to the production frontier of the producer. In the figure the output *y* can be produced by the input *x*. However, with the given input *x*, also a higher scalar output can be produced, $y/\mu*$ which is on the line of the production frontier of the producer (Kumbhakar and Lovell, 2000). Thus, the output distance function equals $D_O(x, y) = \mu^* < 1$.

**Definition 1:** An output distance function equals $D_O(x, y) = \min\{\mu: y/\mu \in P(x)\}$

### 2.1.3 Definition of technical efficiency function

One could use the techniques that have been developed in the estimation of the production frontier (where there is a single-output) to estimate the stochastic output distance function (where there are multiple-outputs) (Kumbhakar and Lovell, 2000). The choice of using multiple-outputs might pose problems, as there is no given choice for a dependent variable. Further, the endogeneity of the regressors could prove the estimation difficult. However, there are several possibilities to estimate the stochastic output distance function if the issues are resolved. In order to view the mathematical form of the stochastic distance function, there has to be some changes of the production frontier model.

The deterministic production frontier model when cross-sectional data of the inputs to produce one output is assumed:

$$y_i = f(x_i; \beta) \cdot TE_i, \tag{2.1}$$

where $y_i$: is the scalar output of producer $i$, $i = 1, 2, ..., I$.
$x_i$: is the vector of $N$ inputs used to produce $y$.
$f(x_i;\beta)$: is the production frontier.

$\beta$: is a vector of the technology parameters that are to be estimated.
$TE_i$: is the output oriented technical efficiency of producer $i$.
The deterministic production frontier model can be rewritten as

$$\mathrm{TE_i} = \frac{y_i}{f(x_i;\beta)}, \tag{2.2}$$

Equation (2.2) describes technical efficiency of producer $i$, $TE_i$, to be equal to the ratio of the observed output, $y_i$, and the maximum feasible output, $f(x_i;\beta)$. The equation further describes that the observed output can achieve a maximum feasible value of output if and only if $TE_i=1$. If $TE_i<1$, it will provide a measurement of the difference between the observed output and the maximum feasible output.

When the production frontier model is deterministic, as it is in equation 2.2, it will ignore random shocks that can affect the producer and thus the output. In order to include the random shocks to the output and producer, the deterministic production frontier is rewritten with the component of exp{vi} and thus one get the stochastic production frontier model:

$$y_i = f(x_i;\beta) \cdot \exp\{v_i\} \cdot \mathrm{TE_i}, \tag{2.3}$$

In equation (2.3), $[f(x_i;\beta) \cdot \exp\{v_i\}]$ is the stochastic production frontier. $f(x_i;\beta)$ is the deterministic part of the stochastic production frontier which is common to all producers and the stochastic part that catches the random shocks, $\exp\{v_i\}$, and is unique for all producers. The stochastic production frontier can be rewritten into:

$$\mathrm{TE_i} = \frac{y_i}{f(x_i;\beta)\cdot\exp\{v_i\}}, \tag{2.4}$$

Equation (2.4) is somewhat similar to equation (2.2), but there is a difference regarding when the producer is technical efficient or not. In equation 2.4, technical efficiency is the ratio of the scalar output of producer $i$ and the maximum feasible output and the random shocks that could occur. However, the producer will achieve the maximum feasible output if and only if $TE_i = 1$. If $TE_i < 1$, it will measure the difference between the observed output and maximum feasible output. The difference will be measured in the environment of the stochastic component, $\exp\{v_i\}$, which is allowed to be different from producer to producer. Technical efficiency can be estimated through either equation (2.1) and (2.2), the deterministic production frontier or equation (2.3) and (2.4), the stochastic production frontier. The latter one is the better as it includes the effects of random shocks in the producer's environment (Kumbhakar and Lovell, 2000).

The single-output stochastic production frontier can be written as:

$$y_i = f(x_i;\beta) \cdot \exp\{v_i - u_i\}, \tag{2.5}$$

Where $TE_i = \exp\{-u_i\}$ and $\exp\{v_i\}$ is the random-noise error component. The single-output stochastic production frontier can be rewritten as follows:

$$\frac{y_i}{f(x_i;\beta)} = \exp\{v_i - u_i\}, \tag{2.6}$$

Equation 2.6 can also be written as:

$$D_O(x_i, y_i; \beta) = \exp\{v_i - u_i\}\ , \tag{2.7}$$

In the single-output case $D_O$(x,y) = y/f(x). This can then be transferred to the multiple-output case, which gives equation (2.7). Equation (2.7) it can be rewritten to a stochastic distance function model:

$$1 = D_O(x_i, y_i; \beta) \cdot \exp\{v_i - u_i\}\ , \tag{2.8}$$

Where $E(u_i|u_i\text{-}v_i) \geq 0$, $\exp\{u_i\text{-}v_i\} \geq 1$ and $D_O\ (x_i,y_i;\beta) \leq 1$. These properties will give the inverted measure of the technical efficiency. Equation (2.8) has to be converted in order to be an estimable regression model. Distance functions are linearly homogenous in outputs $D_O(x_i,\lambda y_i;\beta) = \lambda D_O(x_i,y_i;\beta)$, $\lambda > 0$. The property of linearly homogenous in outputs makes it easier to convert the equation to an estimable regression model. However, there are two ways to proceed. Kumbhakar and Lovell (2000) normalize the output vector with the inversion of the Euclidean norm by setting $\lambda = |y_i|^{-1} = (\sum_m y^2_{mi})$. This normalization generates that $D_O(x_i,y_i/|y_i|;\beta) = |y_i|^{-1}\cdot D_O(x_i,y_i;\beta)$ which in turn gives $D_O(x_i,y_i;\beta) = |y_i|\cdot D_O(x_i,y_i/|y_i|;\beta)$. This can be substituted into equation (2.8) and when dividing both sides with $|y_i|$ one get an estimable regression model:

$$|y_i|^{-1} = D_O\left(x_i, \frac{y_i}{|y_i|}; \beta\right) \cdot \exp\{u_i - v_i\}, \tag{2.9}$$

As mentioned, there are different ways in order to normalize the output vector. Emvalomatis *et al.* (2008) have from Coelli and Perelman (1996) used a logarithmic approach instead in order to normalize the output vector. Their estimable regression model is:

$$-\log y_{it}^M = \log D_O\left(x_{it}, \frac{y_{it}}{y_{it}^M}; \beta\right) + u_{it} + v_{it}\ , \tag{2.10:a}$$

The difference between equation (2.9) and (2.10:a) is what kind of dataset that is used. Kumbhakar and Lovell have given a regression model that is appropriate for cross-sectional data while the model in equation (2.10:a) is used for panel data. However, the equation could be rewritten so that the notation of time, *t*, is excluded and thus the regression model could be used for cross-sectional data:

$$-\log y_i^M = \log D_O\left(x_i, \frac{y_i}{y_i^M}; \beta\right) + u_i + v_i\ , \tag{2.10:b}$$

The model in equation (2.10:b) can be estimated by maximum likelihood, but distributional assumptions has to be placed on both $u_i$ and $v_i$. If $v_i$ represents a random noise in the production process, it is assumed to have a normal distribution (Emvalomatis *et al.*, 2008). There is however difficulties to assign a distribution to $u_i$. Aigner *et al.* (1977) assumed it to be half-normal distributed, Stevensson (1980) assumed it to be truncated-normal distributed and Greene (1990) assumed it to be either exponential or gamma distributed. Emvalomatis *et* *al.* (2008) choose to base the distribution of the $u_i$ by firm-specific characteristics, $z_i$, which is based on the work by Battese and Coelli (1995).

### 2.1.4 Distributions

There are different distributions one could assign the error component $u_i$ and the stochastic noise component $v_i$. The different distributions are the normal-half normal model, the normal-exponential model, the normal-truncated normal model and the normal-gamma model (Kumbhakar and Lovell, 2000).

The distributional assumption for the normal-half normal model according to Kumbhakar and Lovell (2000) is the following one:

I. $v_i \sim \text{iid } \mathrm{N}(0, \sigma_v^2)$
II. $u_i \sim \text{iid } \mathrm{N}^+(0, \sigma_u^2)$, meaning nonnegative half-normal.
III. $v_i$ and $u_i$ are distributed independently of each other and of the regressors.

The first distributional assumption, I, is the same for all of the distributions. The third assumption, III, is also the same for every distributional model. The second assumption, II, is unique for every distribution. The second assumption for the normal-exponential model is $u_i \sim \text{iid exponential}$ and for the normal-truncated normal model it is $u_i \sim \text{iid } \mathrm{N}^+(\mu, \sigma_u^2)$ and for the normal-gamma model it is $u_i \sim \text{iid gamma}$.

When the econometric estimation is conducted in TSP, two parameters will be assigned different values in order to give the model its appropriate distribution. The distribution parameters are $\sigma$ and $\lambda$ and are estimated together with the technology parameters ($\beta_i$'s) (Kumbhakar and Lovell, 2000). The explanation of $\sigma$ and $\lambda$ are shown in equation (2.11) and (2.12) below.

$$\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2} , \tag{2.11}$$

$$\lambda = \sigma_u / \sigma_v , \tag{2.12}$$

Where $\lambda$ indicates how much $\sigma_u$ and $\sigma_v$ contributes to the error component ($\varepsilon$) in the econometric model (Kumbhakar and Lovell, 2000). The values of $\sigma$ and $\lambda$ should both be positive. As $\lambda \to 0$ either $\sigma_v^2 \to +\infty$ or $\sigma_u^2 \to 0$, this implies that the model being estimated will be an OLS production function with no inefficiency. As $\lambda \to +\infty$ either $\sigma_v^2 \to 0$ or $\sigma_u^2 \to +\infty$, this implies that the model estimated will be a deterministic production function with no "noise".

### 2.1.5 Functional form

The most commonly used functional forms when estimating the technical efficiency are the Cobb-Douglas- and the translog function (Bravo-Ureta *et al.*, 2007). The stochastic frontier production function is defined as (Battese, 1992):

$$Y_i = f(x_i; \beta)\exp(V_i - U_i), \tag{2.13}$$

The functional form is specified for $f(x_i; \beta)$ that could be either Cobb-Douglas or translog (Battese, 1992). The original Cobb-Douglas functional form is defined as:

$$Y_i = AX_1^{\beta_1}X_2^{\beta_2}, \tag{2.14}$$

Where *Y* equals the output for firm *i*, which is produced by the inputs $X_1$ and $X_2$ and where $\beta_1$ and $\beta_2$ can be interpreted as elasticities of output with respect to $X_1$ and $X_2$ respectively. In order to get a linear function of the Cobb-Douglas, the natural logarithm can be introduced. The Cobb-Douglas functional form as a regression model is defined as (Battese, 1997):

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + V_i, \tag{2.15}$$

In the above Cobb-Douglas model, firm *i* is assumed to produce one output $Y_i$ with two inputs $X_{1i}$ and $X_{2i}$. The $V_i$ for the firms are assumed to be uncorrelated random errors (Battese, 1997). The natural logarithm is widely used in regression modeling and another functional form can be defined as (Bogetoft and Otto, 2011):

$$\ln\big(f(X)\big) = \beta_0 + \sum_{i=1}^{m} \beta_i \ln X_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij} \ln X_i \ln X_j \tag{2.16}$$

Equation (2.16) is a production function that is quadratic in the logarithms of its arguments and is usually called a translog function. The *A* in the function correspond to a matrix, which is assumed symmetric (Bogetoft and Otto, 2011). In order for the matrix to be symmetric the condition that $A_{ij} = A_{ji}$ for all $A_{ij}$ for $i \neq j$ must be satisfied.

There are implications to the Cobb-Douglas functional form (Bogetoft and Otto, 2011). The first order derivative of equation (2.15) give betas that are constant and the second derivative equals zero. This is partly why the translog functional form is favored as the first order derivative give a linear function and the second order derivative give constants. Another feature of the translog function is that the second order derivative of the estimates determine how the inputs and the outputs interact (Bogetoft and Otto, 2011). With the assumption of homogeneity, the data can determine if the inputs and the outputs are either substitutes or complements.

## 2.2 Technical inefficiency

When the technical efficiency has been assessed for all firms, that is $U_i$, the technical efficiency score can be econometrically estimated against exogenous parameters that affect the technical efficiency either positively or negatively and which could be interpreted as sources of inefficiency. Three models that could be used in order to estimate the parameters that contribute to technical inefficiency (Battese and Broca, 1997). Battese and Coelli (1992) proposed the first model. The first model is presented in equation (2.17) below.

$$U_{it} = \{\exp[-\eta(t - T)]\}U_i \tag{2.17}$$

$\eta$ is an unknown parameter that needs to be estimated. $U_i$ corresponds to random variables that are iid distributed and non-negative (Battese and Coelli, 1992). The model is designed for panel data, and will thus not be used in this study.

Battese and Coelli (1995) proposed the second model. The second model is presented in equation (2.18) below.

$$U_{it} = z_{it}\delta + W_{it} \tag{2.18}$$

$z_{it}$ is a vector of explanatory variables that are associated with technical inefficiency (Battese and Coelli, 1995). $\delta$ is a vector of unknown parameters that will be estimated and $W_{it}$ correspond to unknown random variables. $W_{it}$ is assumed to be independently distributed with a zero mean and unknown variance. It is obtained by a truncation of the normal distribution. Huang and Liu (1994) proposed the third model. The third model is presented in equation (2.19) below.

$$U_{it} = z_{it}\delta + z_{it}^{*}\delta^{*} + W_{it} \tag{2.19}$$

$z_{it}^{*}$ is a vector of values that interact appropriately with the variables in $z_{it}$ and $x_{it}$ (the inputs in the production function) and $\delta^{*}$ is a vector of unknown parameters to be estimated (Huang and Liu, 1994). The third model is a non-neutral stochastic frontier as the inefficiency effects are closely related to the input variables in the production function. When using a translog production function in order to estimate the stochastic frontier, the function for the inefficiency effects could be extensive, depending on how many inputs that are used and how many explanatory variables will be included for the estimation of the technical inefficiency.

## 2.3 Previous applications of SFA

Bravo-Ureta *et al.* (2007) made a meta-regression analysis of 167 studies that included farm level technical efficiencies. The farms are both located in developing as well as developed countries and the technical efficiency scores were calculated by stochastic frontiers models, non-parametric and parametric deterministic models. The different agricultural products in the studies were rice, maize, other grains, other crops, dairy and cattle, other animal and whole farm production. Bravo-Ureta *et al.* (2007) conclude that the largest average mean of technical efficiency is found for other animals (84.5 %). It had however the least number of observed studies that calculated the technical efficiency scores. Out of the 167 studies, 117 used the stochastic frontier as the method to calculate the technical efficiency scores.

In table 1 below, the average mean technical efficiency (AMTE) for stochastic and deterministic methodology is presented. The majority of the studies have relied on parametric models, using panel data, the Cobb-Douglas functional form and a primal representation of the technology (Bravo-Ureta *et al.*, 2007).

Bravo-Ureta *et al.* (2007) conclude that the average mean technical efficiency is greater for the stochastic models than the deterministic models. They further discuss the application of different functional forms used, and find that for deterministic models the average mean technical efficiency is greater for the Cobb-Douglas functional form than for the translog. However, when using stochastic models, the values for the functional forms are the opposite.

This is however not statistically significant according to the meta-regression analysis made by Bravo-Ureta *et al.* (2007).

*Table 1. Average mean technical efficiency for stochastic and deterministic methodology (Bravo-Ureta et al., 2007:64).*

| Category | Deterministic | | | Stochastic | | | AMTE |
|---|---|---|---|---|---|---|---|
| | Avg. | Max. | Min. | Avg. | Max. | Min. | |
| *Approach* | | | | | | | |
| Parametric | 70.2 | 95.5 | 26.0 | 77.3 | 100.0 | 17.0 | 76.3 |
| Non-parametric | 78.3 | 100.0 | 35.0 | - | - | - | 78.3 |
| *Data* | | | | | | | |
| Panel | 77.5 | 96.0 | 35.0 | 78.4 | 96.0 | 43.0 | 78.2 |
| Cross-sectional | 72.8 | 100.0 | 26.0 | 75.2 | 100.0 | 17.0 | 74.2 |
| *Functional form* | | | | | | | |
| Translog | 72.6 | 95.5 | 41.0 | 76.3 | 100.0 | 17.0 | 75.7 |
| Cobb-Douglas | 68.1 | 77.6 | 49.0 | 79.7 | 99.8 | 24.0 | 78.9 |
| Others | 64.6 | 79.7 | 26.0 | 83.2 | 86.4 | 66.2 | 68.3 |

Note: AMTE is the average mean technical efficiency of the 167 studies analyzed by Bravo-Ureta *et al.*, (2007).

Several studies point to the limitation of the Cobb-Douglas functional form (Battese, 1992; Bravo-Ureta and Pinheiro, 1993). There are however other studies concluding that the functional form has a small impact on the estimated technical efficiency scores (Koop and Smith, 1980). In one study made by Ahmad and Bravo-Ureta (1996), the Cobb-Douglas functional form was rejected for a simplified translog form; they however concluded that the technical efficiency measures where little affected by the choice of functional form.

## 2.4 Returns to scale

From the literature, it is concluded that there should be homogeneity for the variables in production function used in the analysis of the stochastic frontier. The homogeneity is related to the returns to scale of the inputs used in order produce one or more outputs. When constant returns to scale is evident, the unit increase in inputs increase the output by the same unit. If output increases less or more, there is increasing or decreasing returns to scale evident, respectively, in the production function. The assumption of homogeneity or constant returns to scale does not always hold in previous studies.

*Table 2. Studies that contains test of constant returns to scale of the production function. Source: studies analyzed by Bravo-Ureta et al. (2007). Note: own summary and adaptation.*

| Product studied | Complete no. of studies | Studies not found | No test of CRS | Test of CRS |
|---|---|---|---|---|
| Whole farm | 10 | 1 | 3 | 6 |
| Crops | 13 | 6 | 6 | 1 |
| Wheat | 4 | 1 | 2 | 1 |
| Other crops | 12 | 2 | 4 | 6 |
| Grains | 2 | 1 | 0 | 1 |
| Other grains | 8 | 0 | 3 | 5 |
| Sum | 49 | 11 | 18 | 20 |

Table 2 present studies that have been included in the meta-regression analysis conducted by Bravo-Ureta *et al*. (2007). Out of the 117 studies analyzed by Bravo-Ureta *et al*. 49 studies are of interest due to their connection to crop production. Out of the 49 studies, 11 studies could not be found or read. 18 studies did not either mention constant returns to scale or conduct a test for constant returns to scale. 20 out of 49 studies test for or mention constant returns to scale. The 20 studies that mention constant, increasing or decreasing returns to scale are presented in table 3 below.

Even though the assumption that constant returns to scale holds for the production function that is used in the stochastic frontier analysis, it can be concluded from table 3 that this is not always the case. In 5 out of 20 studies, there is constant returns to scale present in the production functions analyzed. In the majority of the studies, there is decreasing returns to scale present in the production function and in three studies, there is increasing returns to scale present. According to Schmidt (1977), the economies of scale can have a dampening effect on technical inefficiency when the functional form used is a Cobb-Douglas production function.

*Table 3. Studies that mention constant returns to scale or test for it. Note: own summary and adaptation.*

| Author | Product | Country | Comment |
|---|---|---|---|
| Bagi *et al*. (1982) | Whole farm | USA | Both reject and accept CRS[4] |
| Brümmer *et al*. (2000) | Whole farm | Germany | CRS not rejected |
| Brümmer *et al*. (2001) | Whole farm | Slovenia | CRS not rejected |
| Huang *et al*. (1984) | Whole farm | India | DRS evident |
| Rezitis *et al*. (2002) | Whole farm | Greece | DRS, statistically significant |
| Rezitis *et al*. (2003) | Whole farm | Greece | DRS, statistically significant |
| Bravo-Ureta *et al*. (1997) | Crops | Dominican Rep. | CRS rejected |
| Binam *et al*. (2004) | Other crops | Cameroon | CRS rejected |
| Coelli *et al*. (2004) | Other crops | Papua New Guinea | IRS evident |
| Heshmati *et al*. (1997) | Other crops | Sweden | IRS evident |
| Lansink (2000) | Other crops | Netherlands | From IRS to DRS[5] |
| Tzouvelekas *et al*. (2001) | Other crops | Greece | CRS rejected |
| Tzouvelekas *et al*. (2001) | Other crops | Greece | Both reject and accept CRS[6] |
| Giannakas *et al*. (2000) | Wheat | Greece | CRS rejected |
| Ivaldi *et al*. (1994) | Grains | France | DRS, statistically significant |
| Battese *et al*. (1993) | Other grains | India | DRS, not certain[7] |
| Battese *et al*. (1997) | Other grains | Pakistan | DRS, not statistically significant |
| Hadri *et al*. (2003) | Other grains | England | Roughly CRS, not tested[8] |
| Hadri *et al*. (2003) | Other grains | England | IRS and DRS, statistically significant |
| Kurkalova *et al*. (2003) | Other grains | Ukraine | DRS evident |

Note: CRS - constant returns to scale, IRS - increasing returns to scale and DRS - decreasing returns to scale.

[4] Bagi *et al*. (1982) analyze both organic and conventional production and there is thus two different production functions present.

[5] Lansink (2000) use data for Netherlands over time and conclude that that the elasticities of the input variables change over time, from summing to greater than one to summing to less than one.

[6] Tzouvelekas *et al*. (2001) analyze both organic and conventional production and there is thus two different production functions present.

[7] In Battese *et al*. (1993) the elasticities of the input variables sum to less than one but there is no hypothesis test conducted whether the null hypothesis should be rejected or accepted.

[8] In Hadri *et al*. (2003) the elasticities of the input variables sums close to unity but there is no hypothesis test conducted whether the null hypothesis should be rejected or accepted.

## 2.5 Summary

Stochastic frontier analysis is a method used in order to measure the efficiency of a firm or an organization that uses one or multiple inputs to produce one or multiple outputs. There are cost, profit- and technical efficiency, where Farrell introduced the econometric modeling of technical efficiency in 1957. The technical efficiency of a firm equals the ratio between the actual output given its inputs and the frontier output, being the output that the firm could produce if the right technology were available.

This study will use stochastic frontier analysis in order to analyze the performance of Swedish farms engaged in horse related activities. There are those that use the deterministic production function in order to analyze the technical efficiency of firms, but it is not used when firm produce two or more outputs. Instead, the stochastic frontier approach allows for statistical noise, which implies that favorable or unfavorable conditions for the firms are captured. The statistical noise can be used in order to estimate the factors that contribute to technical efficiency. There are three models that could be used and in this study the model proposed by Battese and Coelli (1995) will be used.

The most commonly used functional forms when estimating the technical efficiency are the Cobb-Douglas- and the translog function. This study will use the Cobb-Douglas production function for the estimation of the technical efficiency of Swedish farms. The Cobb-Douglas production function is the functional form that is generally used in the studies studied by Bravo-Ureta *et al.* (2007). The authors also conclude that the stochastic model is preferred to the deterministic model.

# 3 Data

In this chapter, the data from the survey is discussed and statistically analyzed. The variables that are used in the estimation of the technical efficiency of the farms is presented and the survey data used in this study is compared to data originating from a similar survey made by the Swedish Board of Agriculture in 2011.

## 3.1 Survey

The survey used in this study was conducted, supervised and implemented by LRF Konsult[9]. The department of economics at the Swedish university of agriculture made the design of the survey. The survey is a source of information for a 3-year research project that is implemented by the department of Economics at the Swedish university of Agriculture and the Norwegian department of Agricultural research. The purpose of the project is to learn more about the horse sector and its actors, both on the producer side as well as on the consumer side. The data collected by the survey will be part of the basis of the research project, where the survey will act as a source of information regarding farmers that is in some part connected to the horse sector. The connection is horse activities in different ways, but the survey is mainly focused on horse boarding.

LRF Konsult have a register with 2200 farms that are involved in horse boarding activities. Out of these 2200 farms, the questionnaire was sent to 1000 of them. Of the 1000 questionnaires 428 were returned, which yields a response rate of about 43%. The answers to the surveys have all been compiled in an Excel sheet. The questions in the survey cover demography, education, distance to nearest city, what type of business the farm have, what type of production, capital, labor, what type of horse activities, income and costs as well as management and marketing. The complete questionnaire is presented in appendix A.

## 3.2 Analysis of data

In this sub-section, there will be an analysis of the data in order to exclude potential outliers and observations that contain incomplete or strange data. There is also an explanation of how the variables have in some cases been developed and calculated, followed by a statistical analysis of these variables. The variables of importance for the assessment of technical efficiency are capital, labor, land, output and intermediate consumption.

### 3.2.1 Incomplete data and outliers

Before calculating technical efficiency scores, the first task was to undertake a preliminary analysis of the data and exclude all observations that were not complete. In the survey, three questions are of great importance for analyzing the farms as well as needed in order to calculate the technical efficiency. The first question is number 12 on the first page in the survey, which give information about the production activities on the farm. The second and third questions (numbered 39 and 40, respectively) give information about total income and

[9] LRF is The Federation of Swedish Farmers, where LRF Konsult is a subsidiary working as advisers to e.g. farmers.

total variable costs of the sampled farms. If there were no information given in either question 39 or 40, the observed farm was excluded from the dataset. Taking into account these incomplete answers to questions 39 and 40 results in the exclusion of 288 observations. Furthermore, 44 of the 428 questionnaires have a missing front page where question 12 is found. In some cases the front page were never included in the survey when sent to the farms and in other cases they were missing when collected. This results in the exclusion of 44 observations.

Out of the 288 observations excluded due to incomplete answers regarding income and costs, about 65 % of the questionnaires have no information about income or costs. About 35 % have some information about the income and costs of the farm, but are despite this incomplete. Even if there is some explanation to why 44 observations are excluded, there is no explanation to why 288 respondents has not given the complete information regarding their accounting other than that they did not reveal such information or the one responding to the survey did not know such information.

The measurement of the farms capital is very important. For this purpose, capital is divided into several items including machinery, buildings, horses owned by the farm and their market values. For 11 observations, there is no information about the machinery or the buildings on the farm. The observations are thus excluded as the market value of the machinery and the buildings cannot be calculated. Another eight observations are excluded as the agricultural production on the farm is related to other cattle than horses (e.g. beef or pork production). In the survey, there is a question regarding the number of horses owned by the farm and the market value of them, but there is no question regarding the number of other cattle on the farm or the market value of them. If there had been information regarding the number and what type, then the market value could have been calculated. With the exclusion of incomplete observations regarding capital, the number of observations remaining in the dataset equals to 77.

A major problem that might arise when dealing with data is the problem of observations that differ very much from the other sample observations, namely the so-called outliers. They are sometimes easily observed in the data but might sometimes be recognized too late and will impose problems during regression modeling. The method used to detect these outliers is to develop various scatter diagrams between key economic variables such as total income and total variable costs.

Two scatter diagrams presented in Appendix B are used to identify the outliers. In the first diagram, it is obvious that there is an outlier as the observation has an income more than five times larger than the second largest (50 million SEK contra 10 million SEK). After the exclusion of this outlier, there are still two outliers remaining that needs to be excluded. The first one is in the second diagram in appendix B, where there is still one obvious outlier. The second is found when taking the difference between the income and costs, and it show a farm that has its costs exceeding its income with 1.1 million SEK. Further analysis of the data give away four more observations with strange data regarding capital and are also excluded. With these four observations, there are seven observations that are either outliers or that have strange information. Furthermore, there are 12 observations excluded as they have no information regarding agricultural production (this is discussed in depth under subchapter 3.2.2 Variables). The successive steps adopted in the preliminary analysis of the data led us to consider a much-reduced sample of 58 observations that can be used in the empirical estimation of a stochastic frontier production function. For a detailed account of these successive steps concerning the data analysis, see table 4.

*Table 4. Descriptive table of excluded observations. (Own adaptation).*

| | |
|---|---|
| Total number of observations | 428 |
| Observations with no front page | 44 |
| Observations with incomplete info about income and costs | 288 |
| Observations with incomplete info about capital | 19 |
| Observations that are outliers and have strange data | 7 |
| Observations with no specification of agricultural production | 12 |
| Final number of observations used in the calculation of technical efficiency | 58 |

### 3.2.2 Variables

The variables that are of importance for estimating the technical efficiency are the following ones: capital, labor, land, income and intermediate consumption. The information in the survey about these variables differs from each other. The amount of land in hectares is given in the survey, as well as the number of horses owned by the farmer and the market value of the horses. These variables are the only ones that can be taken directly from the survey and then be used in the econometric work for estimating the technical efficiency. The following paragraphs will briefly explain how the remaining variables of importance for the calculation of technical efficiency have been measured so that they could get proper values. These variables concerns capital, labor, income and intermediate consumption.

The first variable of concern is capital, which is obtained by separately assessing the values of machinery and buildings. The information given by the respondents concerning machinery is the year of fabrication, the amount of horsepower's (if a tractor, loader or combine), and the width of the machine (if a combine, seeder, harrow, plow, spray or baler). The information given regarding the buildings is the size in square meters while storage for grain and vegetables is in cubic meters. The calculation of the market value for machinery is shown in equation (3.1) below.

$$\text{Market value}_{\text{Machinery}} = \text{RVP} \cdot \text{Market price}_{\text{Machinery(2012)}} \tag{3.1}$$

where $RVP$ is the remaining percentage value.

The market price of some of the machinery is found in Databoken 2013 (Internet, Agriwise, 2013) where the database from Agriwise has enough information about the machinery that can be combined with the information given from the survey. For the plow and spray, the market price is found in Lantmannen redskap (2005). The market prices in the catalog were for 2004. The prices are thus recalculated in order to represent the market prices of 2012. The market value of the machinery in 2012 is then calculated by multiplying the remaining value percentage, $RVP$[10], with the market price.

The market values of the buildings on the farms are found by assessing the taxation value of the buildings. This is not the real market value, but may be the closest approximation of it. Multiplying N with Fe finds the taxation value. Both $N$ and $F_e$ depend on the location of the farm. There are three measures for the location of the farm, rural (G), peri-urban (M), and close to city (T). These three different locations are given in the survey for each observation. The value of $F_e$ is specifically dependent on the size of the building and the location. The specific value for each building is found in the appendix of Skatteverket (2010) which is a handbook for assessing the taxation values of farm buildings. The $N$ value is an average value

[10] Further explanation of RVP and the function of it is found in appendix C.

for Sweden[11] depending on the location (G, M or T) of the farm. The taxation value, R[12], is then calculated by multiplying $N$ and $F_e$ and the equation is shown below.

$$R = N_{G,M,T} \cdot F_e \tag{3.2}$$

where $N$ is the location of the farm, either rural (G), peri-urban (M) or closer to city (T). and $F_e$ is the location of the farm and the size of the building.

The information from the survey regarding income is the total income for the farm and the income from horse activities. As is shown in figure 6, the income from the horse activities comes from three different sources (boarding, grazing or other horse activities) but in some cases, it could come from only one source.

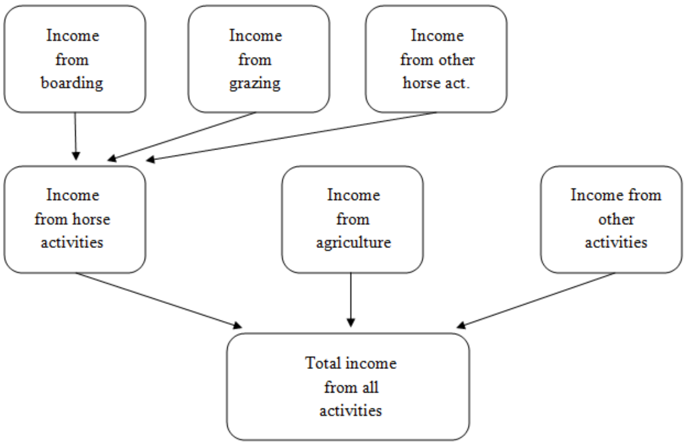



*Figure 6. The income distribution of the farms. Note: own adaptation.*

With the information about the total income from horse activities, it is easy to estimate the income from the other sources being either agricultural production or other production activities, where the agricultural production consists of crop production and grazing activities for horses.

Table 5 below presents the different types of production activities on the farms. The farm could have more than only one of either of the production activities. On the left hand side of the table, the other production processes are presented in italics. On the right hand side of the table, the different types of horse activities are presented. Agriculture is put in italics, as its activity is already included in the types of production on the farm. The answers to the questionnaire regarding farms that has an agricultural production is rather inconsistent. From the table below it is concluded from question 12[13] that 24 farms have an agricultural production and from question 19,[14] it is concluded that there are 26. In question 21[15] in the survey, there are those that have answered that their main source of income is agriculture, but they have not stated what type of agricultural production is conducted. The final number of farmers that have an agricultural production is 58.

[11] The $N$ values are in fact a given value for each municipality in Sweden. However, as there is no information in the survey regarding where the farms are located in Sweden, the average value has been calculated.
[12] Further explanation of $R$ is found in appendix D.
[13] "What production activities are conducted on the farm?"
[14] "What type of horse activities are included in your enterprise?"
[15] "Is the horse business your main source of income? If no, what is?"

*Table 5. Production activities on the farms. Source: (Boarding data). Note: own adaptation.*

| Production on farm | | Type of horse activities | |
|---|---|---|---|
| Grain | 18 | Boarding | 57 |
| Beef | 0 | Trotting | 2 |
| Poultry | 1 | Riding school | 1 |
| Pork | 0 | Health and recreation | 0 |
| Forage | 6 | Goods and services | 9 |
| Dairy | 0 | Trail riding | 0 |
| *Contracting* | 14 | Tourism | 2 |
| *Forestry* | 28 | Breeding | 13 |
| *Gardening* | 0 | Other | 0 |
| *Estate management* | 7 | *Agriculture* | *26* |
| *Other services* | 4 | | |

The following equations explain how the income from agriculture and other production activities are derived. The information that is given in the survey is found on the left hand side in equation (3.3). When subtracting the total income with the income from horses, there will be income left, which will be distributed to either income from agricultural production or income from other production activities such as contracting, forestry, estate management and other services.

$$\text{Income}_{\text{Agr,Other}} = \text{Income}_{\text{Tot}} - \text{Income}_{\text{Horse}} \tag{3.3}$$

The income that is left is then multiplied with either the share of agricultural production or share of other production activities as shown in equation (3.4) and (3.5). The share is calculated by adding up all the agricultural production activities as well as other production activities for each individual firm, and dividing it with the total number of production activities.

$$\text{Income}_{\text{Agr}} = \text{Income}_{\text{Agr,Other}} \cdot \text{Share of agricultural production} \tag{3.4}$$

$$\text{Income}_{\text{other}} = \text{Income}_{\text{Agr,Other}} \cdot \text{Share of other production} \tag{3.5}$$

As there is no information regarding what production activities that is the main source of income (except for question 21 that is taken into account when calculating the share) this is assumed the best way of calculating how the income is distributed.

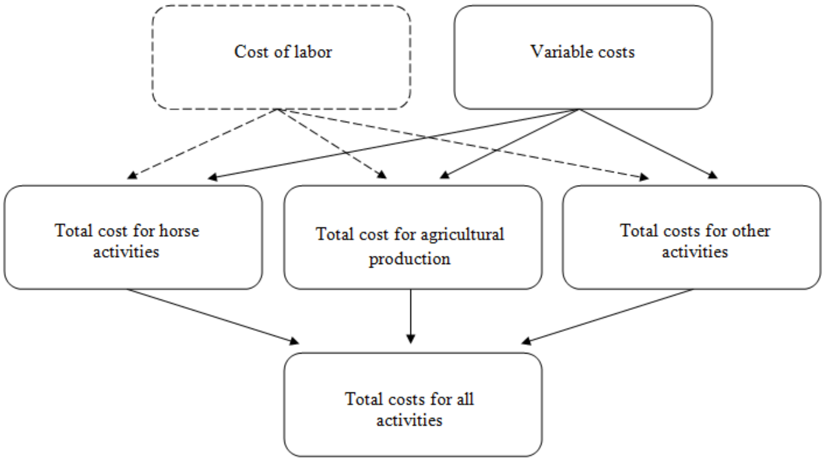



*Figure 7. The sources for the variable costs of the farms and the distribution of them. Note: own adaptation.*

The costs are calculated in a similar way that of the incomes, but there are however some differences. The distribution of the costs for the farms is presented in figure 7 above. Other variable costs are assumed to be the costs of intermediate consumption, e.g. feed for the horses, fuel for tractors and more. The cost of labor is connected to the different total costs of the farm as it is assumed included in some observations. In 18 observations there are hired labor on the farm and the average costs for the hired labor is calculated with the help of the average salary from Databoken 2013 by Agriwise). It is assumed that the cost of labor is not included in the total cost for the farm if the cost of labor exceeds the total cost. If the cost of labor is less than the total cost, it might be included in the total cost. The cost of labor exceeds the total costs in 10 out of 18 observations, which implies that for those 10 observations the cost of labor is not included in the total costs. Thus in 8 observations the cost of labor is included in the total costs and needs to be excluded from it in order to assess the variable labor, which is the value of labor measured in labor cost. If this is not done, the cost of labor might be included in both the variable of labor and intermediate consumption (being the total cost for the farm).

The explanation of how the cost of labor is calculated is shown in equation (3.6). Hired labor and the given fulltime equivalent (FTE) of the farms is found in the survey. When the fulltime equivalent exceeds one, there is hired labor. 23 501 SEK is the average salary per month found in Agriwise, and all is multiplied with 12 months in order to give the total cost of labor for one year.

$$\mathrm{Cost}_{\mathrm{Labor}} = \text{Hired labor} \cdot 23\,501\ \text{SEK} \cdot 12\ \text{months} \qquad (3.6)$$

The cost of labor is then excluded from the 8 observations it is included. This is done in order to find the value of labor and use it as a separate variable in the calculation of the technical efficiency. The value of labor is then simply the fulltime equivalent multiplied with the average salary for a year. This calculation is shown in equation (3.7).

$$\mathrm{Value}_{\mathrm{Labor}} = \mathrm{FTE} \cdot 23\,501\ \text{SEK} \cdot 12\ \text{months} \qquad (3.7)$$

where *FTE* is the fulltime equivalent.

As there is no information regarding in which production activity the labor is used, the labor cost will be excluded from the total costs of the farm and then the variable costs allocated for horse activities and agricultural production are recalculated. The calculation for the new total cost for horse activities is shown in equation (3.8). The share of total cost is the share that the total costs of horse activities has of the total cost of the farm.

$$\mathrm{Cost}_{\mathrm{Horse}} = \mathrm{Cost}_{\mathrm{Tot}} - \mathrm{Cost}_{\mathrm{labor}} \cdot \text{Share of } \mathrm{Cost}_{\mathrm{Tot}} \tag{3.8}$$

When the new total cost of horse activities have been calculated, the remaining total costs will be distributed to either agricultural production or other production activities. The calculation is found in equation (3.9).

$$\mathrm{Cost}_{\mathrm{Agr,Other}} = \mathrm{Cost}_{\mathrm{Tot}} - \mathrm{Cost}_{\mathrm{Horse}} \tag{3.9}$$

Now with the remainder of the total costs calculated (being the summed total cost of agricultural production and other activities). The total cost for the agricultural production and other production activities are calculated by equations (3.10) and (3.11). The share of agricultural and other production is the same share used when calculating the income distribution.

$$\mathrm{Cost}_{\mathrm{Agr}} = \mathrm{Cost}_{\mathrm{Agr,Other}} \cdot \text{Share of agricultural production} \tag{3.10}$$

$$\mathrm{Cost}_{\mathrm{Other}} = \mathrm{Cost}_{\mathrm{Agr,Other}} \cdot \text{Share of other production} \tag{3.11}$$

### 3.2.3 Statistical analysis of data and variables

In this subchapter, there will be a statistical analysis of the variables that are used when calculating the technical efficiency. The previous subchapter only described how the different values of the variables were derived; in this subchapter, the values of the calculations will be presented.

In table 6 below, the statistical analysis of labor is presented. It is shown that there is a small amount of labor on the farms, with the mean of fulltime equivalent equaling 0.95 and the mean of number of workers being 1.49. This implies that there are more workers on the farms than there are fulltime work opportunities. Many farmers having other jobs beside their farm activities can explain the mean of the fulltime equivalent. In the Boarding data, it is concluded that 31 (or 42 %) of the farmers has stated that their main source of income is coming from off farm activities (mainly other occupation). 16 out of 70 farms has hired labor, and the mean of the hired labor is 1.02. The amount of hired labor, given in fulltime equivalent, is as low as 0.1 and 0.25 and in other cases as high as 2.

*Table 6. Statistical analysis of labor. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | Mean | Min | Max | Std. dev. | Observations |
|---|---|---|---|---|---|---|
| FTE | Fulltime equivalent | 0.95 | 0.05 | 3.00 | 0.74 | 70 |
| LABC | Value of labor | 268 063 | 14 100 | 846 039 | 208 159 | 70 |
| LW | Number of workers | 1.49 | 0.00 | 6.00 | 1.41 | 66 |

In table 7 below, the quartiles of the labor is presented. The 1st quartile corresponds to the 25 % of the observations and its values, the 2nd 50 %, the 3rd 75 % and the 4th 100 %. There is a possibility to show the skewness of the data through the help of quartiles. Analyzing the fulltime equivalent it is shown in table 3.4 that 75 % of the values are less than or equal to 1. In the last 25 % of the values, they differ and are larger than 1. The same percentage is found when examining the value of labor. This is because the value of labor is calculated by the fulltime equivalent. The number of workers are at the maximum 2 in 75 % of the values. This relates very much with the hired labor that has a mean of 1.02, as 1 person working on the farm is assumed to be the owner and any labor above 1 will correspond to hired labor.

*Table 7. Quartile analysis of labor. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | 1st quartile | 2nd quartile | 3rd quartile | 4th quartile |
|---|---|---|---|---|---|
| FTE | Fulltime equivalent | 0.44 | 1.00 | 1.00 | 3.00 |
| LABC | Value of labor (SEK) | 124 262 | 282 013 | 282 013 | 846 039 |
| LW | Number of workers | 0.00 | 1.50 | 2.00 | 6.00 |

In table 8 below, the statistical analysis of land is presented. The three variables presented are arable land, grazing land and forest where their corresponding values are given in hectares. The mean of the arable land of the farms are quite low with only 27.62 hectares. The farms with the most hectares have 175. The mean of the grazing land is even smaller but where the hectares of forest are the largest. It might seem conspicuous that the amount of grazing land is small when all the farms have horse activities, but it is concluded from the Boarding data that only 0.7 % of the total number of horses on the farms are actually grazing. The rest are either housed in stables or other.

*Table 8. Statistical analysis of land. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | Mean | Min | Max | Std. dev. | Observations |
|---|---|---|---|---|---|---|
| AL | Arable land (ha) | 27.62 | 0.00 | 175 | 34.85 | 69 |
| AG | Grazing land (ha) | 9.96 | 0.00 | 95 | 14.69 | 69 |
| AF | Forest (ha) | 48.09 | 0.00 | 300 | 77.56 | 70 |

In table 9 below, the quartiles of land is presented. It is shown in the table that the amount of arable and grazing land for the farms is small, which has been concluded from the previous table. 75 % of the farms arable land is below 33 hectares, for grazing it is 10 hectares. For forest on the other hand, it is 49. In the three cases, there is a large "jump" in the last 25 % of the values.

*Table 9. Quartile analysis of land. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | 1st quartile | 2nd quartile | 3rd quartile | 4th quartile |
|---|---|---|---|---|---|
| AL | Arable land (ha) | 7 | 15 | 33 | 175 |
| AG | Grazing land (ha) | 2.3 | 5 | 10 | 95 |
| AF | Forest (ha) | 0 | 16 | 49 | 300 |

In table 10 the statistical analysis of the machinery on the farms are presented. The most interesting aspect in the table is the number of observations, that implies the number of farms that actually has any machinery (and thus capital) on the farm. The most farms has one tractor, but it decreases rapidly when asked if they have two or more. When it comes to the other machinery, there are no equipment that is found more often on the farms than other, except for harrow and plow. In the survey, 33 farms have grain production, then why there are only 11 farm that has a combine and 14 that has a sower. This is explained by the fact that many farms, 25 (or 34 %) rent machinery or equipment either as a service or as it is. The most rented equipment is the sower, where there are 15 farms renting it either as a service or as it is. Other machinery or equipment that is commonly borrowed are combines, plows, harrows and plant sprays. The summation of the machinery give the information that there are 65 farms that own machinery or equipment and this variable is the used as part of the capital for the farms.

*Table 10. Statistical analysis of machinery in SEK. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | Mean | Min | Max | Std. dev. | Observations |
|---|---|---|---|---|---|---|
| KMT1 | Tractor 1 | 154 288 | 1 275 | 744 464 | 183 471 | 62 |
| KMT2 | Tractor 2 | 101 351 | 1 173 | 569 003 | 153 246 | 33 |
| KMT3 | Tractor 3 | 71 389 | 1 935 | 605 036 | 151 021 | 16 |
| KMT4 | Tractor 4 | 319 581 | 2 225 | 949 660 | 446 573 | 4 |
| KMLO | Loader | 48 666 | 6 077 | 219 216 | 75 529 | 9 |
| KMCO | Combine | 307 519 | 7 350 | 1 422 508 | 415 982 | 11 |
| KMSO | Sower | 19 219 | 3 341 | 68 773 | 20 751 | 13 |
| KMHA | Harrow | 17 271 | 2 179 | 64 213 | 19 139 | 18 |
| KMLI | Lister | 27 059 | 2 487 | 168 065 | 48 199 | 11 |
| KMPL | Plow | 24 289 | 1 168 | 122 904 | 28 021 | 18 |
| KMSP | Plant spray | 11 763 | 1 196 | 33 967 | 12 978 | 5 |
| KMBA | Baler | 5 585 | 3 575 | 8 947 | 2 019 | 7 |
| KMA | Sum of machinery | 305 248 | 4 098 | 2 344 962 | 472 118 | 65 |

In table 11 below, the statistical analysis of the buildings is presented. From the table it is evident that stables and barns are the most common buildings on the farms. The amount of general farm buildings is low, but it could be explained by the fact that farms with either beef or pork production has been excluded. In some cases, there are observations with incomplete data. In 11 cases, there are incomplete data regarding the size of the building in $m^2$ and thus the taxation value has not been calculated. These observation with incomplete data has however not been excluded, as they have in other cases stated the $m^2$. There will however be a shortage of their true capital in buildings in the end. Even though there are missing information in 11 cases, all farms has stated at least one building, which is presented in the bottom right corner of table 11.

*Table 11. Statistical analysis of buildings in SEK. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | Mean | Min | Max | Std. dev. | Observations |
|---|---|---|---|---|---|---|
| KBFB | Farm building | 256 773 | 228 602 | 587 574 | 156 513 | 5 |
| KBST | Stable | 446 239 | 115 805 | 1 842 385 | 370 884 | 68 |
| KBRH | Riding hall | 1 603 142 | 547 941 | 2 981 525 | 694 413 | 23 |
| KBBA | Barn | 257 734 | 55 019 | 1 814 797 | 282 570 | 55 |
| KBMH | Machine hall | 139 428 | 46 555 | 561 820 | 109 235 | 38 |
| KBWO | Workshop | 110 216 | 46 555 | 300 283 | 60 064 | 38 |
| KBSG | Grain storage | 371 087 | 189 042 | 774 925 | 199 028 | 16 |
| KBSV | Vegetable storage | 209 848 | 209 848 | 209 848 | 0 | 1 |
| KBU | Sum of buildings | 1 441 339 | 283 846 | 5 002 934 | 1 216 780 | 70 |

In table 12 below, the quartiles of the capital is presented. In the table, it is shown that the values for the machinery and the buildings are skewed to the left. In the case of machinery, 75 % of the population has less than 400 000 SEK worth of machinery and only 25 % of the population has machinery worth of more than 400 000 SEK. In the case of buildings, 50 % of the population has buildings worth 2 000 000 SEK and the last 25 % has buildings worth up to 5 000 000 SEK. The small numbers of the machinery come from farms that have machinery that is very old combined with the fact that they rent either the equipment as it is or the service of it.

*Table 12. Quartile analysis of capital in SEK. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | 1st quartile | 2nd quartile | 3rd quartile | 4th quartile |
|---|---|---|---|---|---|
| KMA | Sum of machinery | 35 888 | 142 041 | 380 114 | 2 344 962 |
| KBU | Sum of buildings | 536 699 | 1 010 422 | 2 017 456 | 5 002 934 |

In figure 8 below the agricultural output is plotted against the arable land of the farms. It is shown that a large amount of the farms has little arable land and little agricultural output. 63 out of the 70 farms have some sort of agricultural production on their farm, but the output of the agricultural production is the main source for only 26 farms. The observations, which have a fair amount of agricultural output and, a considerable amount of arable land are those farms that are producing grain.

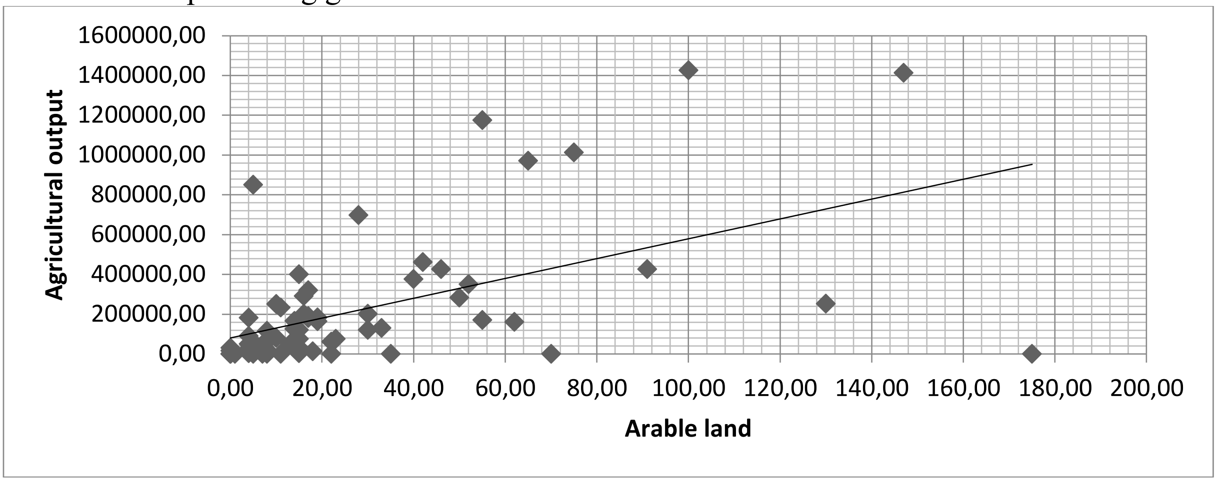



*Figure 8. Agricultural output in SEK vs. arable land in ha. Source: (Boarding data). Note: own adaptation).*

In figure 9 below the total output is plotted against the total capital of all the farms. The spread and the difference of the farms are a little bit more obvious in this figure than figure 8, but still there are many farms that are smaller and are clustered together while there are in some cases other farms that either own a large amount of capital but with a relatively small total output or vice versa. It should be noted that the total output for the farms are never zero in any observation, but could be around 10 000 SEK and is thus in that case close to zero in the diagram. A preliminary thought about the plotted observations is that the observations found below the line could be less efficient as they use a large amount of capital in their production of a small amount of total output. However, the largest part of the total value of capital is the value of buildings. Thus, it would be more representative to plot the agricultural output against machinery and horse output against the buildings (as the highest values of the separate buildings come from buildings that are related to horse activities).

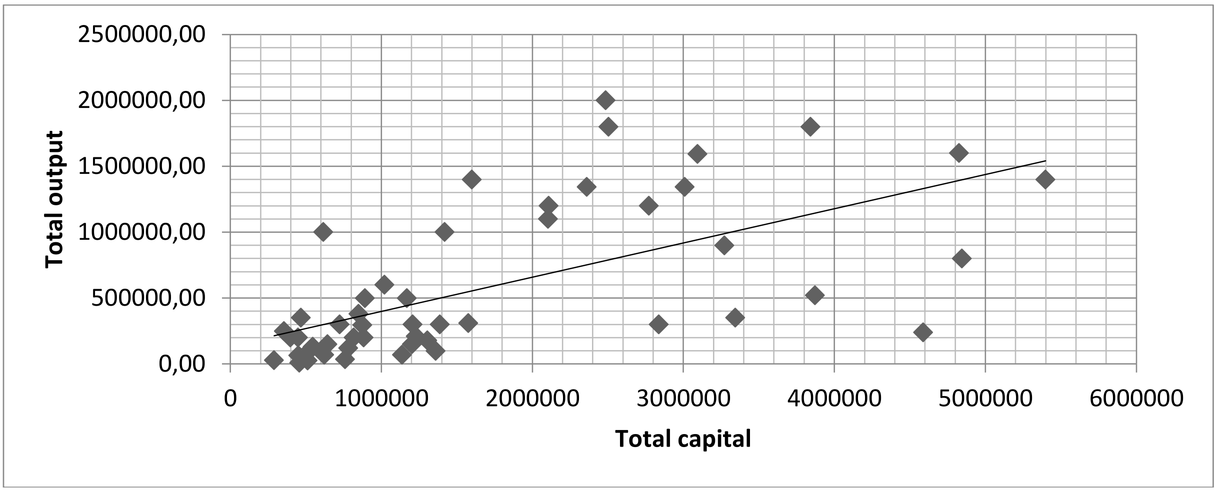



*Figure 9. Total output vs. total capital in SEK. Source: (Boarding data). Note: own adaptation.*

In figure 10 below the agricultural output and the machinery owned are plotted against each other for all 70 observations. One fact that remain and that is obvious in the figure below is that most farms are clustered around a small agricultural output and a low value of the machinery owned by the farm. What differs between the values in figure 9 and 10 below is that there are zero values for agricultural output. Out of the 70 farms, 12 farms have no agricultural production. 8 out of the 12 observations have a value for other output on the farm while the remaining 4 observations only have horse related output.

As previously mentioned regarding the large amount of capital used and the low output produced, it is clear that there are two farms that has a low value of the agricultural output in respect to the value of the capital being used with respect to other farms. These two farms have however output from other farm activities corresponding to the same value as the agricultural output. The machinery could thus be assumed to be used in the other activities on the farm and not only in the agricultural production (as well as for horse related activities).

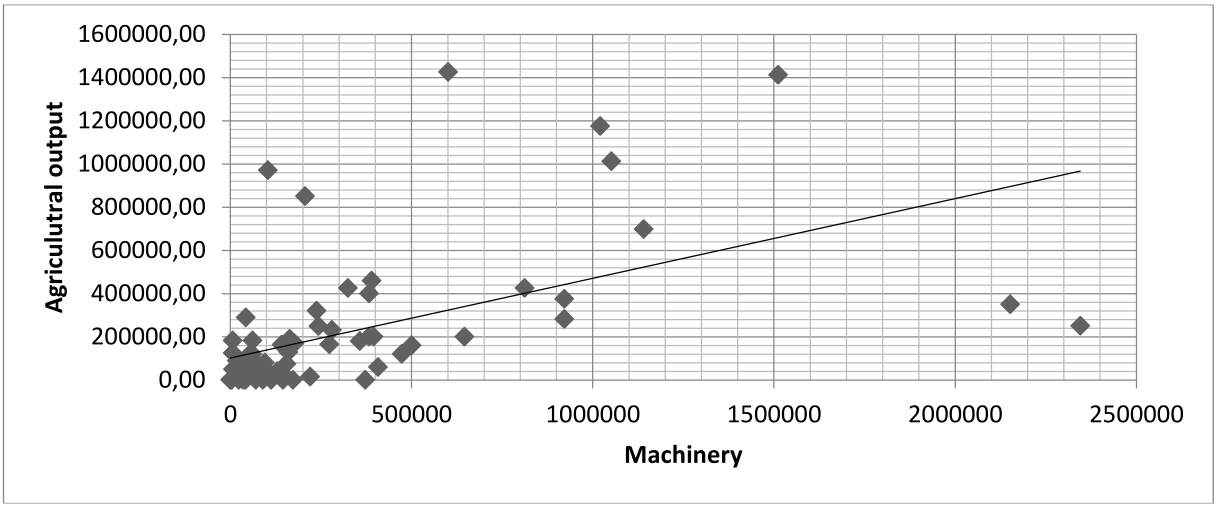



*Figure 10. Agricultural output vs. machinery in SEK. Source: (Boarding data). Note: own adaptation).*

In figure 11 below the output of horse activities is plotted against the total value of buildings. The spread of the observations are not as large as in figure 10, but as in all figures, the main number of observations is clustered in the bottom left hand corner. It should be noted that there are no zero values for the output of horse activities.

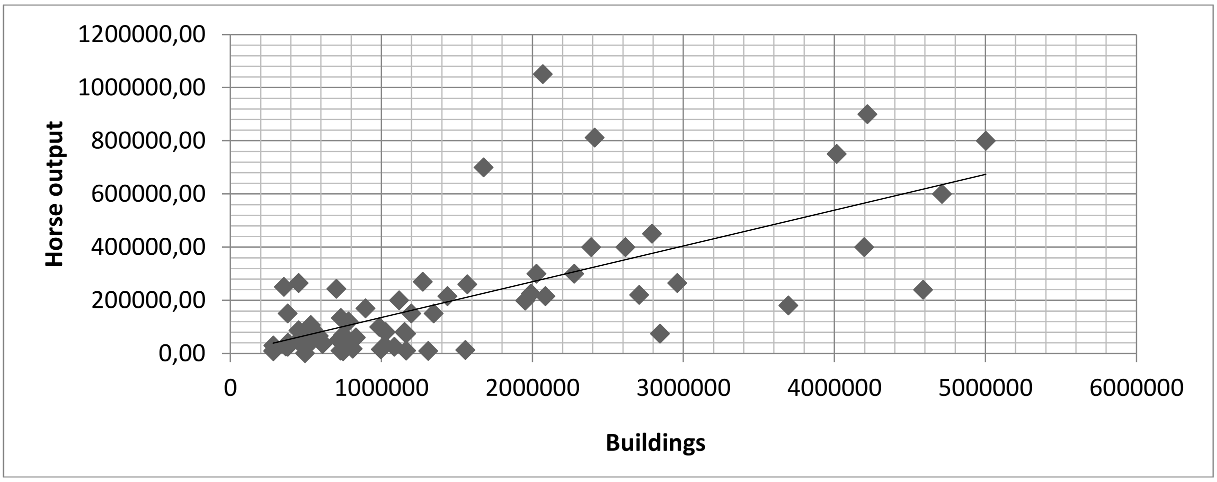



*Figure 11. Horse output vs. buildings in SEK. Source: (Boarding data). Note: own adaptation.*

In table 13 below it is concluded that there are no zero values for the output from horse activities on the farms. It is also concluded that the output from horse activities has the output in value terms compared to the other outputs. The production activity that has the largest output is that of agriculture with 1 425 000 SEK and a mean of 216 009 SEK, even though it has the largest output the production process has the second largest spread (shown by the standard deviation). Considering the costs of the different production activities, the costs for agricultural production is the highest while the costs for the horses are the least.

*Table 13. Statistical analysis of output and costs in SEK. Source: (Boarding data). Note: own adaptation.*

| Var. name | Var. definition | Mean | Min | Max | Std. dev. | Observations |
|---|---|---|---|---|---|---|
| IAG | Agri. output | 216 009 | 0.00 | 1 425 000 | 324 114 | 73 |
| IOT | Other output | 174 817 | 0.00 | 1 387 500 | 321 712 | 73 |
| IHO | Horse output | 194 267 | 1 000 | 1 050 000 | 235 932 | 73 |
| TCAG | Agri. costs | 177 246 | 0.00 | 1 107 584 | 222 807 | 73 |
| TCOT | Other costs | 136 659 | 0.00 | 1 100 000 | 220 008 | 73 |
| TCHO | Horse costs | 112 582 | 0.00 | 900 000 | 158 669 | 73 |

As been mentioned previously, 12 farms have no agricultural production. Furthermore, 26 farms have no other production on the farm beside horses and agriculture. The aim of this study is to evaluate the technical efficiency of two production processes on the farm, namely horse related activities and agricultural production. This implies that the information about the costs and output from other on farm production will be excluded in the calculation of the technical efficiency. Out of the 70 observations, 58 observations or farms have a value for the output from horse related activities as well as a value for the output from agricultural production.

When looking at the diagrams in figure 8 to 11 it is evident that there might be a problem with heteroscedasticity[16] as the many of the farms with less output are clustered together and the farms with greater output are scattered around in the figures. Heteroscedasticity could arise when there are outliers present in the data (Gujarati and Porter, 2009). However, the outlier that has been found in the data has been excluded as well as other strange data that could provoke heteroscedasticity. Another source for heteroscedasticity is skewness in one or more

[16] Heteroscedasticity is when the variance of the disturbances ($u_i$) for the variables in a regression model are different from each other $E(u_i^2) = \sigma_i^2$, thus meaning that they are not constant (Gujarati and Porter, 2009).

of the variables included in the econometric model (Gujarati and Porter, 2009). From the quartile analysis of the variables, it is evident that there is a skewness for smaller values of the variables. Even though there being a skewness of the variables, it is interesting to calculate the technical efficiency in order to verify if there is a difference in the efficiency between a small-scale and a large-scale farm.

Heteroscedasticity might pose several problems, especially when using ordinary least squares in the regression model (as the variances are not estimated the appropriate way) the analysis of the *t*- and *F*-tests could be wrong (Gujarati and Porter, 2009). However, even though heteroscedasticity is proven to be observed in some cases and it could be a bad reason to exclude an otherwise good formulated regression model (Fox, 1997).

## 3.3 A previous survey

The Swedish board of agriculture did in 2011 a large survey of the horse sector in Sweden. The survey had eight pages containing 35 multiple-choice questions. Overall, they sent 2 775 surveys to farms that have both horses and riding schools. 1 820 was the total number of answers they got back which gives a response rate of about 66 %.

The purpose of the survey was to get a picture of the horse sector in Sweden, who is the typical horse owner, and for what purpose the horse is used. The questions in the survey were targeted in order to get information about what feeding is given to the horses, are other horses in contact with the owners' own horses, the amount of land the horses use and is able to use, what the horses are used for, how the horse business looks like and the owners' view of their ability to develop or increase their equine-related business in the future.

From the survey, it is concluded that 60 % of the horse owners in Sweden are women (Jordbruksverket, 2012). However, there is a large difference between men and women concerning what sort of work they do in the horse sector. The majority of men are trainers in trotting and galloping while women either work for or own riding schools and businesses related to horse tourism.

According to the survey and the data collected by Jordbruksverket, almost half of the sample has no education related to horses (47 %). The other half that does have an education related to horses are educated in or as trainer/instructor (20 %), other (18 %), forage (15 %) and riding instructor (13 %). Other education implies that the person either has been self-thought (7 %) or taken separate courses about horses (8 %). The education related to horses is presented in figure 12 below.

Even if half of the sample has no education related to horses, the sample is highly educated elsewhere. 47 % of the sample has an education that is higher than high-school level and where 30 % has been educated at a university level.

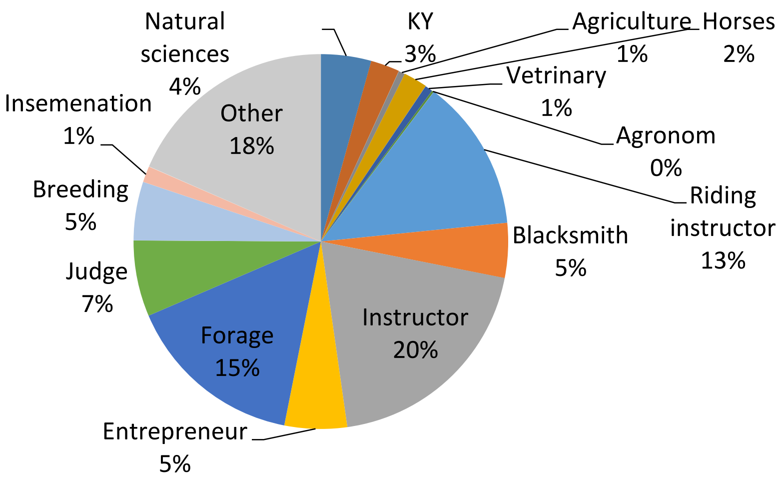



*Figure 12. Education related to horses. Source: Jordbruksverket. Note: own adaptation and KY is vocational education.*

The third important question regarding the characteristics of the typical Swedish horse owner was what type of business he or she has. In the data collected by Jordbruksverket, it is concluded that there are two types of horse owners. The first is the owner that has horses for a hobby or has a hobby business (56 %) while the second have horses for business purposes, which implies that the business is a main source of income (53 %). From the data, it is concluded that 9 % of the respondents have both horses, as a hobby and hobby business and thus the hobby could be their main source of income. Another explanation is that the Swedish taxation office might label the business as a hobby business even if the owner might have it as a main source of income. The use of horses owned for hobby reasons is presented in figure 13 below. The horses are mainly used for riding (49 %), but also grazing care (17 %), exercise or as labor (16 %) and breeding (10 %).

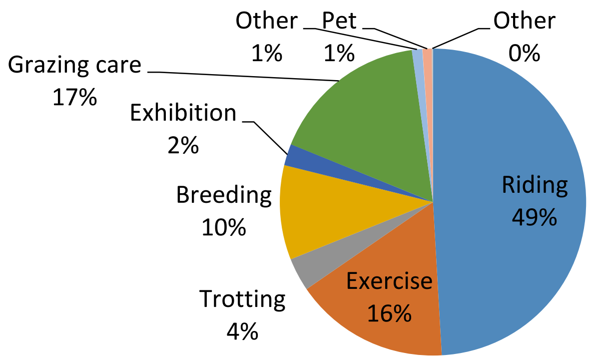



*Figure 13. For what purpose the hobby owners use their horses. Source: Jordbruksverket. Note: own adaptation.*

From the data provided by the survey from Jordbruksverket, it is concluded that there are 3 245 workers employed in the horse different horse activities of the observed sample. The average number of workers employed for each horse related activity in the sample is 1.86, but it is shown by statistical analysis that 50 % of the observed sample has no workers employed at all. Between 50 and 75 % of the observed sample has less than two workers employed, while the last 25 % has between two and 112 workers employed. This implies that the distribution of workers employed in the observed sample is very much skewed towards zero. The same holds for the analysis of the fulltime equivalent for the observed sample where 75

% of the population has a fulltime equivalent less than 1. The remaining 25 % of the observed sample has a value of the fulltime equivalent between 1 and 28.

Jordbruksverket ended their survey by asking a question if the business is planning to hire more workers in the two forthcoming years. Only 51 % of the observed sample answered the question and 42.7 % answered "no" to that question while 8.4 % answered "yes". The 8.4 % correspond to 146 observations answering "yes" to the question, 4 out of the 146 observations do not have any workers employed. The majority of the observations answering yes have more than two workers employed, meaning that it is the larger business' that are planning to hire more workers in the future.

There is also another question in the survey by Jordbruksverket regarding the future of the horse establishments. The question is whether horse owners believe that they will develop their business in the two forthcoming years. The majority of the observed sample answered "no" to that question (66 %). However, out of the 34 % that answered yes, there are between 12 and 20 % that answer that they will not only develop one part of their business but two or more. In figure 14, the different development plans are presented. Even if the majority of the horse owners do not plan for any development in their business or plan to hire workers in the future, the number of horses do still increase in Sweden.

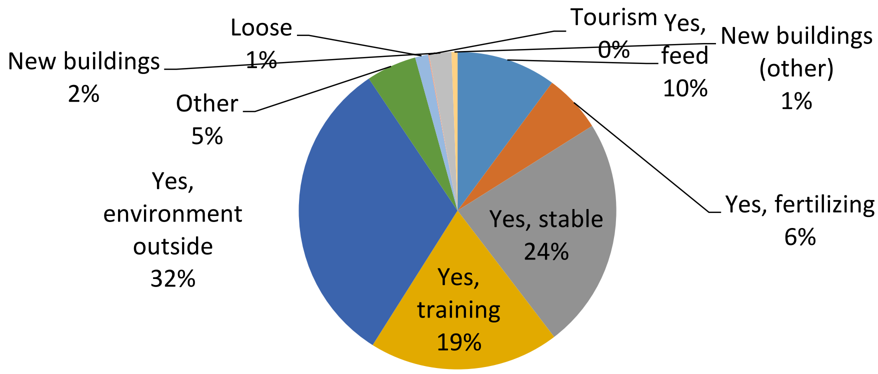



*Figure 14. Plans for development of horse business in the coming two years. Source: Jordbruksverket. Note: own adaptation.*

In the survey made by Jordbruksverket there are also two questions regarding other types of businesses that the horse owner is involved in. The first question is regarding what type of other business that is related to horses and the second question refers to what type of business that is not related to horses. The other type of business' that horse owners are involved in are boarding (18 %), riding school (16 %) and breeding (15 %). There are several other businesses related to horses and all 14 of them are presented in figure 16 below. Agriculture and forestry (5 %) is the 7$^{th}$ type of other business related to horses that are conducted by the horse owners. Even though the percentage is fairly low, there are a fair amount of horse owners that produce their own forage (see figure 15 below). 45 % of the horse owners substitute their total demand for forage by 1 % or more from their own production. 31 % of the horse owners produce enough forage to feed their horses; they have thus no need to buy forage from any other producer. Despite this, the majority (55 %) of the observed sample relies on forage that is produced in the agricultural sector.

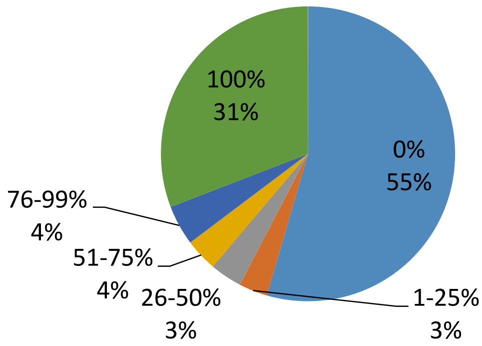



*Figure 15. Percentage substituted by own production of forage. Source: Jordbruksverket. Note: own adaptation.*

The second question regarding business' that are not related to horses but are conducted by the horse owners are presented in figure 16 below. Out of the observed sample, only 37 % answered that they had no other business that are not related to horses, while 10 % answered that they do. The percentage presented in figure 17 below are the share of the 10 % of the observed sample that answered "yes" where it is concluded that the majority have agriculture as another business not related to horses (42 %).

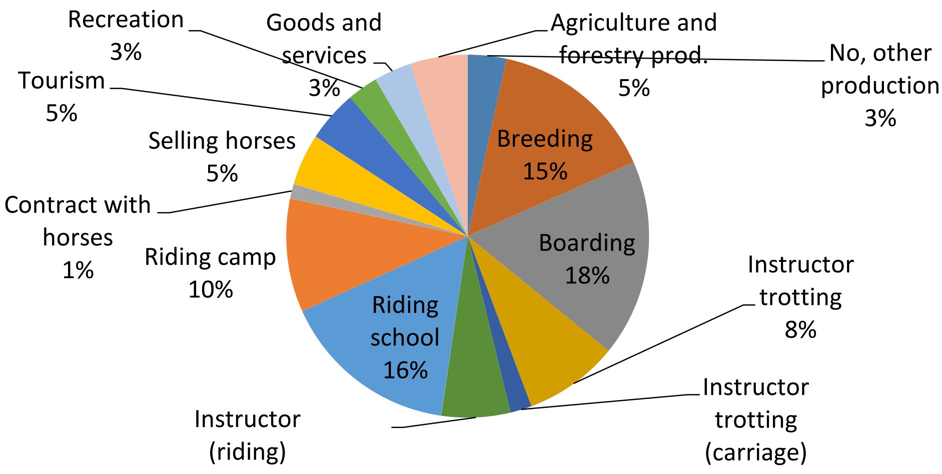



*Figure 16. Other business related to horses. Source: Jordbruksverket. Note: own adaptation.*

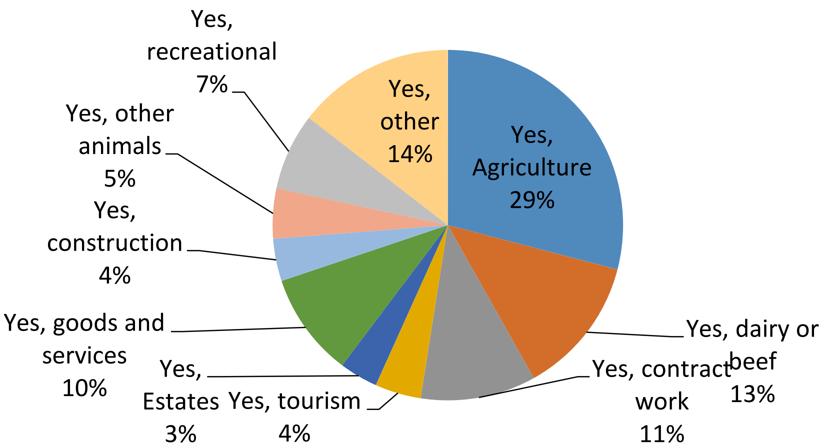



*Figure 17. Business not related to horses. Source: (Jordbruksverket). Note: own adaptation.*

### 3.3.1 Comparison between surveys

In the answers from the survey made by Jordbruksverket (SJV) it is concluded that 40 % of the horse owners were men while 60 % were women (Jordbruksverket, 2012). In the survey made by SLU, it is the other way around where approximately 25 % of the horse owners are women and 75 % are men. However, there is no difference between the two surveys concerning the age of the horse owners. In SJV:s survey, 60 % of the horse owners are the age between 41 and 65 years old and in the survey made by SLU the average age is 56 years old.

There is no large difference between the two surveys with respect to education. Both of the surveys have similar questions regarding education and the percentage of the answers are similar. It can also be concluded that almost half of the observed sample, 47 %, has no education related to horses. In the case of the survey conducted by SLU, there are 55 % answering that they do not have an education related to horses.

The question regarding the main source of income for the horse business was not included in the surveys sent to trainers, breeders or riding schools as it was assumed that they had their activity as their main source of income. However, the question was posed in the surveys that were sent to the ordinary horse owners together with the question if it was a hobby business or not. More than 60 % answered that their main source of income originates from other occupation. In the survey conducted by SLU, only 26 % answered that their main source of income was salary from other occupation and 20 % was from agriculture. In the survey made by SLU, there is a question if the farm has agriculture as a main source of income, there is no similar question in the survey made by the Swedish board of agriculture.

In the survey conducted by SJV it is concluded that approximately 70 % of the horses in Sweden are owned by the horse owner and a little more than 20 % of the horses in the observed sample are boarded. As the survey from SLU is focused on horse boarding, it is evident that the percentage of the answers should be more than the answers from the survey conducted by SJV, which it is with between 70 and 78 % of the observed sample answering that they board horses. The number of horses that are on the farms presented in the survey made by SLU are 4 324 out of which 62 % are boarded, 35 % are own horses and 3 % are grazing.

The survey by SJV gives information regarding how many hectares of grazing land that is available for the horse owners. The amount of hectares available for grazing is approximately 376 500 hectares which with 362 700 horses in Sweden imply that there is one hectare of grazing land for every horse in Sweden. In the survey done by SLU it can be concluded that there is two hectares of grazing land for every horse in the observed sample. As the SLU survey also focus on the agricultural perspective, a major part of the land presented in the answers are of arable/tillable land. The total amount of grazing land, both owned and rented is 8 875 hectares while arable land sums to 33 370 hectares including both owned and rented arable land. Despite arable land not being included in the survey made by SJV, it is presented in the report that approximately 40 % of the horse owners are growing their own forage. In the survey conducted by SLU only 7.5 % answer that they grow forage and 35 % produce grains. The small amount of the observed sample producing forage might be explained by it not being an alternative for an answer in the survey.

One of the last questions in the survey made by SJV is if and what type of business the horse owners are conducting besides their horse business. 55 % answer that they have another business beside their horse activities and where the majority answer that it is either agricultural or forestry businesses that they operate.

## 3.4 Summary

In this chapter, the data from the survey by LRF Konsult has been discussed and statistically analyzed. The analysis of data have been made in order to exclude observations that are not a reliable source of information for calculating the variables included in the stochastic production function. There were 1000 surveys sent to farms that are involved in horse related activities. Out of those 1000 surveys, 428 were sent back which is the data sample used in this study. The final number of observations used in this study is 58 due the exclusion of observations having incomplete information regarding income, costs, capital, and agricultural production or being outliers.

The calculation of the variables that are part of the stochastic production frontier are presented in this chapter. The variables of importance are: output (income) from agricultural production and horse related activities, capital, land, value (cost) of labor and intermediate consumption for the agricultural production and the horse related activities. With the farms not only having agricultural production and being involved in horse related activities, it has been of importance to try to isolate the output and intermediate consumption related to other activities conducted on the farm. This is why there has been recalculations of the total income and total costs of the farm.

From the statistical analysis of the variables, it can be concluded that the majority of the farms are relatively small with respect to land and partly capital. The mean, minimum, maximum and standard deviations of the variables of importance are presented in table 14 below.

*Table 14. Statistical analysis of variables of importance, n=58. Source: (Boarding data). Note: own adaptation.*

| Var. name | Mean | Min. | Max. | Std. Dev. |
|---|---|---|---|---|
| Land[17] (ha) | 80 | 4 | 360 | 76.32 |
| Labor[18] (SEK) | 247 673 | 14 100 | 846 039 | 194 611 |
| Capital[19] (SEK) | 1 977 089 | 318 947 | 8 771 901 | 1 711 884 |
| Intermediate consumption (IC)[20] (SEK) | 437 974 | 9 000 | 1 677 859 | 413 243 |
| Output[21] (SEK) | 236 413 | 13 500 | 1 107 584 | 239 148 |

Comparing the mean total output of the farms in this study and the average total output of crop farms in 2012 for Sweden in the FADN-data, it can be concluded that the farms represented in this study are small compared to the average of Sweden. The average total

[17] The sum of forest, grazing- and arable land.
[18] The value of labor or cost of labor.
[19] The summed value of machinery, buildings and horses owned.
[20] The costs derived from agricultural production and horse related activities.
[21] The income derived from agricultural production and horse related activities.

output of crop farms in the FADN-data is 837 314 SEK[22] (96 173 EUR) and excluding the subsidies for the farms the output is equal to 499 073 SEK[23] (57 323EUR).

[22] The average total output of crop farms in 2012 (96 173 EUR) times the average exchange rate of 2012 (8.706 SEK). Source: Internet, European Commission and Skatteverket, 2014-08-26. Note: own calculation.

[23] The average total output of crop farms in 2012 (96 173 EUR) minus the balance of subs and taxes on farms in 2012 (38 850 EUR) times the average exchange rate of 2012 (8.706 SEK). Source: Internet, European Commission and Skatteverket, 2014-08-26. Note: own calculation.

# 4 Econometric results and analysis

In the following chapter, the results from the estimation of the technical efficiency will be presented. The first results were the Cobb-Douglas and the translog production function. The production functions are analyzed first in order to conclude which one will be appropriate for further estimation of the stochastic frontier. The estimated parameters of the Cobb-Douglas production function had appropriate estimates and were statistically significant. The estimated parameters of the translog production function had appropriated estimates for the parameters. Not all estimated parameters were significant (5 out of 15 parameters were significant at a 10 % level or lower). Despite the fact that the estimated parameters of the translog production function were appropriate and the $R^2$ being greater than the Cobb-Douglas model, the technical efficiency scores could not be obtained as lambda was equal to a great positive number and thus implies that the production function is deterministic and the production function cannot be estimated by stochastic frontier analysis.

## 4.1 Cobb-Douglas production function

The variables that are used in the Cobb-Douglas production function are the output from the agricultural production and the horse activities on the farm (Y), the land used in the production (Land), the labor in amount of hours worked per year, the capital used in the production (Cap) and the intermediate consumption for the farms (IC). The Cobb-Douglas production function is defined in equation 4.1 below.

$$\mathrm{Y_i} = \beta_0 + \beta_1\mathrm{logLand} + \beta_2\mathrm{logLabor} + \beta_3\mathrm{logCapital} + \beta_4\mathrm{logIC}, \tag{4.1}$$

The production function is formulated in a Word-file and with the necessary commands in order to get the ordinary least squares of the function as well as other statistics. The result from TSP is presented in table 15 below.

*Table 15. Estimates and statistics of the Cobb-Douglas production function.*

| Parameter | Estimate | Std. error | t-statistic | P-value |
|---|---|---|---|---|
| Intercept | -3.19642 | 1.51478 | -2.11015 | 0.035** |
| Land | 0.169336 | 0.079868 | 2.12020 | 0.034** |
| Labor | 0.337880 | 0.088862 | 3.80231 | 0.000*** |
| Capital | 0.414023 | 0.127785 | 3.24000 | 0.001*** |
| IC | 0.434910 | 0.086248 | 5.04255 | 0.000*** |

N = 58, $R^2$ = 0.702495, adj. $R^2$ = 0.680042, Log likelihood = -50.8817 Note: (*) significant at 10 %, (**) significant at 5 %, (***) significant at 1 %.

In the results from TSP in table 15 above, the elasticities of the total output with respect to each input is presented. As the elasticity of intermediate consumption is the largest it will add the most to the total output, which implies that if the intermediate consumption on the farm is increased by 1 % the total output of the farm would increase by 0.43 %. The second largest is capital, third labor and fourth land.

When considering land, the farms usually have a small amount of arable as well as grazing land from the beginning. Most of the farms produce grain (but not all), and the total output of the agricultural production are in many cases not that large. As shown in figure 14 there are a couple of farms that has a fair amount of arable land that do also have the largest agricultural output in value terms. However, the majority of the farms have a small amount of arable land and an agricultural output close to the mean of the observed population. In addition, when considering grazing land, the majority of the farms is all differentiated against boarding of horses and thus does not always need a large amount of grazing land. This implies that there have to be large increases in the input of land in order to increase the total output. In respect to arable land, it would be very costly and considering grazing land it could be hard to find available land that is close to the farm if most of it is already used.

Labor has the third largest impact on total output. Previous econometric estimations of the Cobb-Douglas production function have shown different estimates for the labor parameter and many have not been either statistically significant or have had the right effect on total output (a negative effect on total output in econometric estimation of the translog production function). However, when using the amount of hours worked per year as a variable that has been derived from the fulltime equivalent, the estimate and effect of labor is of significance.

The second largest effect on total output is that of capital. As mentioned previously, the majority of the farms are differentiated against horse boarding and thus an increase in buildings would increase their capacity of horses and could thus increase the total output of horses. In addition, some of the farms have other horse activities, such as riding schools and other that requires buildings for their day-to-day business. Not all of the farms have a large amount of machinery (and thus not a high total market value of machinery) to be used in their agricultural production and to be added is that the majority of the machinery and equipment being used are in most cases very old. Some farms even rent the machinery or the service of it. If the machinery would be renewed, or owned together with increasing the number of building or the size of the buildings the total output on the farms would increase.

Intermediate consumption has the largest effect on total output. The intermediate consumption on the farms consists of feed for the horses. There are only 6 farms that produce forage and who could use their own production as feed for the horses. Other consumption may include the fuel for the machinery on the farm.

### 4.1.1 Technical efficiency scores

In order to estimate the technical efficiency scores the production frontier has to be estimated. The estimated Cobb-Douglas production frontier and its parameters are presented in table 15 below.

When analyzing the estimates of the parameters in table 16 it is concluded that that both sigma and lambda has a positive estimate. If the value of lambda had been negative or close to zero there would be no technical inefficiency for the farms explaining the variation between the farms. The estimated parameters are all statistically significant except for the intercept and lambda. The interpretation of the p-value of lambda is that it might not be different from zero. Despite this, the estimate of lambda is given after 50 iterations in TSP. The estimate does not change despite changing the starting values of lambda and sigma.

*Table 16. Function evaluation of Cobb-Douglas production frontier.*

| Parameter | Estimate | Std. error | t-statistic | P-value |
|---|---|---|---|---|
| Intercept | -2.02725 | 2.53700 | -0.79907 | 0.424 |
| Land | 0.40629 | 0.19538 | 2.07946 | 0.038** |
| Labor | 0.34205 | 0.12683 | 2.69689 | 0.007*** |
| Capital | 0.17149 | 0.10408 | 1.64754 | 0.099* |
| IC | 0.38143 | 0.08597 | 4.43667 | 0.000*** |
| Sigma | 1.31213 | 0.30323 | 4.32715 | 0.000*** |
| Lambda | 1.38744 | 1.15258 | 1.20377 | 0.229 |

Log likelihood = -50.4182, (*) Significant at 10 %, (**) significant at 5 %, (***) significant at 1 %.

The statistics of the estimated technical efficiency scores are presented in table 17 below. In figure 18, the estimated technical efficiency scores are presented in a manner so it will be easy to portray the distribution of the efficiency scores.

*Table 17. Statistics of technical efficiency scores (Cobb-Douglas).*

| Mean | Std. dev. | Minimum | Maximum | Median |
|---|---|---|---|---|
| 0.69798 | 0.06103 | 0.51608 | 0.83479 | 0.70156 |

Analyzing table 17 above and figure 18 below it is evident that the efficiency scores for the farms are clustered together at the mean technical efficiency. No farm deviates a lot from the other farms. The only farm that does deviate in some way is the farm that is the most efficient of those compared. The second most efficient farm has an efficiency score of 0.80, the most efficient farm has an efficiency score of 0.83, and thus there is a difference of 0.03 units between them, which is the largest difference in efficiency for all farms considered.

Comparing the results with previous studies made, that use stochastic frontier analysis and the Cobb-Douglas production function, it can be concluded that the efficiency scores are less than those presented in Bravo-Ureta *et al.* (2007). According to the meta-regression analysis done by Bravo-Ureta *et al.* (2007), the mean technical efficiency for a stochastic Cobb-Douglas production function is 79.5 %, which is greater than the mean of the efficiency scores in table 17 above. For cross-sectional data, the mean technical efficiency is less than that of panel-data. However, the mean technical efficiency presented in Bravo-Ureta *et al.* (2007) is greater than the mean in table 17. In Bravo-Ureta *et al.* (2007), it is also concluded that the average mean technical efficiency for "other" animals (84.5 %) is the greatest for all the production processes.

The high value of the average mean technical efficiency for "other" animals could be due to there not being so many studies made regarding the corresponding production process. This study regarding horse activities could partly be included under "other" animals. The low mean of technical efficiency in this study compared to that presented in Bravo-Ureta *et al.* (2007) could be explained by the fact that studies in Bravo-Ureta *et al.* are all technical efficiency scores derived from a single-output case. In this study, there are two outputs, both agricultural output and output from horse related activities, which have been aggregated into a total output of the farm for the econometric estimation. The farms in this study might not be as efficient due to them having two or more production processes (output related to off-farm activities) contributing to the total output of the farm.

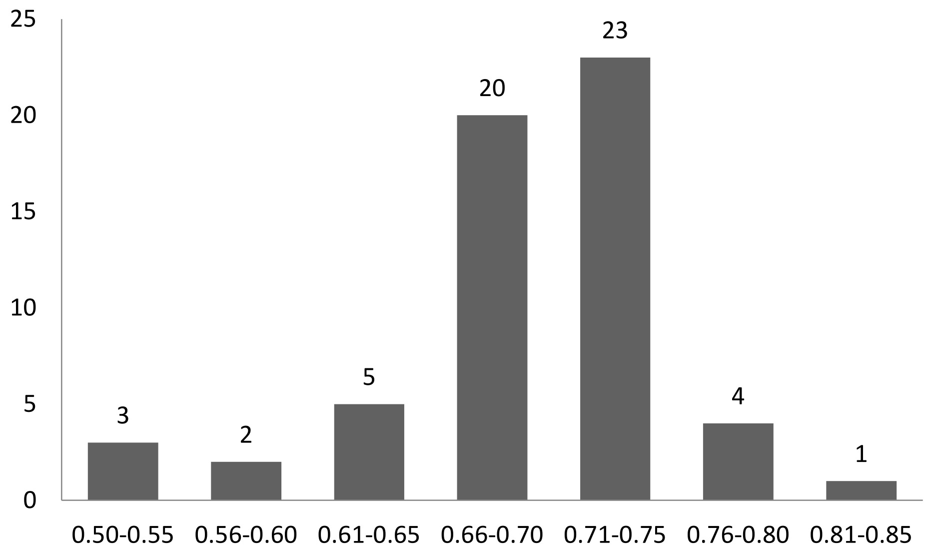



*Figure 18. Distribution of technical efficiency scores (0.50-0.85).*

## 4.1.2 Parameters of technical inefficiency

The estimation of the technical efficiencies was conducted in TSP. However, the estimation of the parameters, which is interpreted as sources of technical inefficiency, were estimated in FRONTIER Version 4.1. The computer program provides maximum likelihood estimation of parameters in both production and cost functions (Coelli, 1996).

The parameters of technical inefficiency can be of different sorts. The parameter can be a dummy variable where it takes either the value of zero or one, e.g. one if the farm is located in a rural area and zero if it is not. The parameter can also be a specific value e.g. the age of the farmer. From the survey there are a few factors that could be used as parameters in the estimation of its effect on the technical efficiency of the farm e.g. education, age, location of the farm, distance to closest city and grants.

This study will focus its attention to the exogenous variables of age, education, location and if the farm receives any grants or not. The analysis of the parameters of technical inefficiency is done in steps where parameters will be excluded from the analysis if they are not statistically significant. When one parameter is excluded, the remaining parameters will be estimated by FRONTIER. In this study, there is an analysis of three different groups of parameters. Equation (4.2), (4.3) and (4.4) below presents these three groups with the *delta*-equations, which is simultaneously regressed with the stochastic production function through maximum likelihood estimation.

$$\mathrm{U_i} = \delta_0 + \delta_1\mathrm{Age}^2 + \delta_2\mathrm{Age} + \delta_3\mathrm{Grant} + \delta_4\mathrm{DisCity} + \delta_5\mathrm{LevelEdu} \tag{4.2}$$

Where $Age^2$ is the age of the farmer squared.
*Age* is the age of the farmer.
*Grant* is a dummy variable where 0: the farm receives no grants and 1: the farm receives grants. The grants are connected to horse related activities.
*DisCity* is the distance to the closest city with inhabitants exceeding 5 000.
*LevelEdu* is the level of education for the farmer taking the value of either 1) secondary school, 2) high school, 3) education after high school and not being 4) university.

$$U_i = \delta_0 + \delta_1 Age^2 + \delta_2 Age + \delta_3 Grant + \delta_4 DisCity \quad (4.3)$$

$$U_i = \delta_0 + \delta_1 Age^2 + \delta_2 Age + \delta_3 Grant \quad (4.4)$$

The result from the maximum likelihood estimation is presented in table 18, 19 and 20 below. The output from FRONTIER gives fewer statistics than from TSP, e.g. the *p-value* of the estimates is not presented in the FRONTIER output. The first five parameters to be estimated are those of the Cobb-Douglas production function (the $\beta's$) and the following six parameters are those of the $z_i\delta_i$ equation. The last two parameters are *sigma* and *gamma* being associated with the variances of $V_i$ and $U_i$.

*Table 18. Statistics of estimated factors of technical inefficiency ($Age^2$, Age, Grant, DisCity and LevelEdu).*

| Parameter | Estimate | Standard-error | T-ratio |
|---|---|---|---|
| Intercept (β) | -1.8995 | 0.6304 | -3.0128 |
| Capital | 0.5182 | 0.1131 | 4.5797 |
| Labor | 0.1101 | 0.0743 | 1.4814 |
| Land | 0.3298 | 0.0857 | 3.8474 |
| IC | 0.4454 | 0.0881 | 5.0534 |
| Intercept (δ) | -0.5514 | 1.0209 | -0.5401 |
| $Age^2$ | -0.0007 | 0.0004 | -1.9608 |
| Age | 0.0688 | 0.0399 | 1.7263 |
| Grant | -0.4196 | 0.1893 | -2.2166 |
| DisCity | -0.0200 | 0.0187 | -1.0677 |
| LevelEdu | -0.2128 | 0.1149 | -1.8522 |
| $\sigma^2$ | 0.0644 | 0.0316 | 2.0388 |
| $\gamma$ | 0.1596 | 0.2487 | 0.6416 |

N= 58, Log likelihood function = 3.1528, LR test of the one-sided error = 11.3207.

The estimation of equation (4.2) is presented in table 18 above. From the estimates of the factors of technical inefficiency it can be concluded that *$Age^2$* increase the technical efficiency while *Age* decrease the technical efficiency of the farms. The estimates of *$Age^2$* and *Age* can, by calculating its first derivatives, present the actual age where the farmer is the most efficient. The farmer will increase its efficiency up to that age and the technical efficiency decreases as the farmer gets older.

The estimate of *Grant* is negative which implies that the technical efficiency will increase if the farm receives any grants. 32 out of the 58 farms receive at least one type of grant. The most farms receive grants from the European Union (28), environmental grants (9), corporate (4), project (3) and leader (1).

The estimate of *DisCity* has a positive impact on the technical efficiency of the farm and thus the further away the farm is from a city of at least 5 000 inhabitants, the more efficient the farm will be. The mean distance to the closest city of at least 5 000 inhabitants is 9.43 km. The estimate of *LevelEdu* also has a positive impact on the technical efficiency of the farm. This implies that the more educated the farmer is, the more efficient the farm will be. 20 farmers have educated themselves at a university level and 21 farmers have studied after their graduation from high school (being other studies than studies at a university level).

Table 19. *Statistics of estimated factors of technical inefficiency (Age², Age, Grant and DisCity).*

| Parameter | Estimate | Standard-error | T-ratio |
|---|---|---|---|
| Intercept (β) | -0.9323 | 0.5665 | -1.6455 |
| Capital | 0.4816 | 0.1096 | 4.3905 |
| Labor | 0.1278 | 0.0722 | 1.7695 |
| Land | 0.2950 | 0.0771 | 3.8278 |
| IC | 0.3508 | 0.0831 | 4.2232 |
| Intercept (δ) | -6.3066 | 7.1388 | -0.8834 |
| $Age^2$ | -0.0027 | 0.0027 | -1.0176 |
| Age | 0.2713 | 0.2792 | 0.9720 |
| Grant | -0.5540 | 0.3298 | -1.6795 |
| DisCtiy | -0.0595 | 0.0477 | -1.2472 |
| $\sigma^2$ | 0.1671 | 0.0709 | 2.3582 |
| $\gamma$ | 0.7813 | 0.1420 | 5.5024 |

N = 58, Log likelihood function = 1.5562, LR test of the one-sided error = 8.1276

The results in table 18 and 19 above show different levels of significance with respect to the t-ratio of the estimates. In table 16, it can be concluded that the estimate of *DisCity* is not greatly significant; this is also the case in table 19. With a further analysis done of the parameters and its estimates it has been concluded that *LevelEdu* and *Age* are collinear and thus the parameter of *LevelEdu* has been excluded in the estimation of the factors of technical inefficiency presented in table 19. With the exclusion of the parameter of *LevelEdu*, there are interesting effects on the estimates of $Age^2$ and *Age* with the estimates having greater effects on the technical efficiency (both negative and positive effects).

Table 20. *Statistics of estimated factors of technical inefficiency ($Age^2$, Age and Grant).*

| Parameter | Estimate | Standard-error | T-ratio |
|---|---|---|---|
| Intercept (β) | -0.7560 | 0.5875 | -1.2869 |
| Capital | 0.4702 | 0.1183 | 3.9749 |
| Labor | 0.1342 | 0.0773 | 1.7351 |
| Land | 0.3006 | 0.0778 | 3.8649 |
| IC | 0.3265 | 0.0873 | 3.7402 |
| Intercept (δ) | -11.2067 | 6.2749 | -1.7859 |
| $Age^2$ | -0.0044 | 0.0023 | -1.9355 |
| Age | 0.4415 | 0.2366 | 1.8663 |
| Grant | -0.8367 | 0.3228 | -2.5919 |
| $\sigma^2$ | 0.3020 | 0.0948 | 3.1855 |
| $\gamma$ | 0.8874 | 0.0604 | 14.6977 |

N = 58, Log likelihood function = 0.1879, LR test of the one-sided error = 5.3911

Further, there is an effect on the t-ratios of the estimates. All of them decrease with the exclusion of *LevelEdu*. With further analysis of the parameters and their estimates, the parameter of *DisCity* is excluded from the estimation due to low significance, which affect the t-ratios of the remaining estimates positively as presented in table 20 above. With just the parameters of $Age^2$, *Age* and *Grant* all estimates are significant at a 5% level or greater, except for the intercept of the stochastic production frontier. The signs of the estimates for the parameters are still the same but the effect on technical efficiency is greater than the previous estimates.

### 4.1.3 Hypothesis testing

It is of interest to analyze the stochastic production function and the parameters of technical inefficiency more closely by examining if the assumption of constant returns to scale holds or not and if the estimates of the parameters of technical inefficiency are different from zero.
The hypotheses being presented in equation (4.5) and (4.6) below are the hypotheses being tested whether there is constant returns to scale present or not in the stochastic production frontier.

$$H_0: \beta_1 + \beta_2 + \beta_3 + \beta_4 = 1 \tag{4.5}$$

$$H_1: \beta_1 + \beta_2 + \beta_3 + \beta_4 > 1 \tag{4.6}$$

The null hypothesis in equation (4.5) will either be accepted or rejected in favor of the alternative hypothesis in equation (4.6). The alternative hypothesis implies that there is increasing returns to scale present in the stochastic production function. The hypothesis will be tested by a t-test, where the null hypothesis will be rejected if the t-value exceeds the critical value of 2.326. The function used in order to obtaining the value of *t* is presented in equation (4.7) below.

$$t = \frac{(R\hat{T}S)-1}{se(R\hat{T}S)} \tag{4.7}$$

Where $R\hat{T}S = \sum_{i=1}^{k} \hat{\beta}_i$ for k inputs,
and $se(R\hat{T}S) = \sqrt{\sum_{i=1}^{k} Var(\hat{\beta}_i) + 2\sum_{i=1}^{k}\sum_{j=1, i\neq j}^{k} Cov(\hat{\beta}_i, \hat{\beta}_j)}$ .
The value of *t* is the difference between the sum of the coefficients of the input variables and the hypothesized value divided by the standard errors of the coefficients of the input variables. The coefficients of the betas are found from the estimation of the production function and the standard errors are found from the covariance matrix of the coefficients. The calculation of the *t* value is presented in equation (4.8) below.

$$t = \frac{1.3561-1}{0.3561} = 2.9376 \tag{4.8}$$

As the *t* value in equation (4.8) exceeds the critical value of 2.326, we reject the null hypothesis of constant return to scale in favor of the alternative hypothesis that increasing returns to scale may be present in the production function.

### 4.1.4 A further analysis of farms

In the following tables, there will be a further analysis of the technical efficiency of the farms. The technical efficiency scores are grouped into four groups and are then compared to different variables, both endogenous and exogenous. It is interesting to see what differentiate the most efficient farms from the least efficient.

In table 21 below, the technical efficiency scores of the 58 farms are compared to their amount of owned land. It can be concluded from the table that the most efficient farms have a small amount of both arable and grazing land.

*Table 21. Analysis of TE scores and land.*

| TE score | N | Arable land (%) | Arable land (ha) | Grazing land (%) | Grazing land (ha) |
|---|---|---|---|---|---|
| 50-59 | 5 | 7.97 | 126 | 23.33 | 122 |
| 60-69 | 21 | 46.65 | 737 | 25.43 | 133 |
| 70-79 | 29 | 43.29 | 684 | 49.90 | 261 |
| 80-89 | 3 | 2.09 | 33 | 1.34 | 7 |

In table 22 below, the technical efficiency scores of the 58 farms are compared to the fact if they are a hobby business or a business operation. It is quite evident that most of the farms see themselves as a business operation and not a hobby business. All three of the most efficient farms see themselves as having a business operation.

*Table 22. Analysis of TE scores and business operation.*

| TE score | N | Hobby (%) | Hobby (N) | Business (%) | Business (N) |
|---|---|---|---|---|---|
| 50-59 | 5 | 1.72 | 1 | 6.90 | 4 |
| 60-69 | 21 | 1.72 | 1 | 32.76 | 19 |
| 70-79 | 29 | 6.90 | 4 | 43.10 | 25 |
| 80-89 | 3 | 0.00 | 0 | 5.17 | 3 |

In table 23 below, the technical efficiency scores of the 58 farms are compared to the output and the capacity of the farms. The output corresponds to the total output that is both agricultural and horse related output. The capacity corresponds to the amount of horses that could be boarded on the farm (as most farms provide such an option as their horse related activity).

*Table 23. Analysis of TE scores output and capacity.*

| TE score | N | Output (%) | Output (SEK) | Capacity (%) | Capacity (N) |
|---|---|---|---|---|---|
| 50-59 | 5 | 9.15 | 2 300 782 | 9.26 | 63 |
| 60-69 | 21 | 49.47 | 12 435 261 | 40.88 | 278 |
| 70-79 | 29 | 40.70 | 10 229 815 | 47.06 | 320 |
| 80-89 | 3 | 0.68 | 170 405 | 2.79 | 19 |

In table 24 below, the technical efficiency scores of the 58 farms are compared to the type of education the farmer have. It is shown that most of the farmers do not have an education related to horses, but many of the farmers have an education in business administration, entrepreneurship and forage.

*Table 24. Analysis of TE scores and education.*

| TE score | N | AgHs | Agr | Train | Entr | BA | For | Judge | Other | NoEdu |
|---|---|---|---|---|---|---|---|---|---|---|
| 50-59 | 5 | 0 | 0 | 1 | 0 | 0 | 2 | 2 | 2 | 3 |
| 60-69 | 21 | 3 | 1 | 1 | 1 | 3 | 5 | 0 | 4 | 10 |
| 70-79 | 29 | 6 | 0 | 2 | 5 | 7 | 6 | 0 | 1 | 15 |
| 80-89 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |

Note: AgHs: agricultural high school, Agr: agronomist, Train: trainer, Entr: entrepreneur, BA: business administration, For: forage, Judge: judge in horse related competitions, NoEdu: no education related to horses.

From the further analysis of the technical efficiency scores and both the endogenous and exogenous variables it could be concluded that the most efficient farms in the sample observation are the smaller farms (with respect to land and output).

# 5 Discussion

The following chapter will discuss chapter 3 and chapter 4. The discussion of chapter 3 will be regarding the data used and how it could have been if the data would have been more extensive. The discussion regarding chapter 4 follow the discussion of chapter 3 and a discussion regarding the technical efficiency and the factors of technical inefficiency.

## 5.1 Data

Out of the 428 answers from the questionnaire, only 58 observations were used in the econometric estimation of the technical efficiency scores and the factors of technical inefficiency. The low number of observations was mainly due to the respondents not answering the questions that were of greatest importance. The question that was of greatest importance and the one the respondents answered the least was the question regarding the total income of the farm, income related to horse activities, the total variable costs of the farm and variable costs related to horse activities. One explanation to why there is low answer rate to question 39 and 40 might be that question 37 is formulated in a way that the farmer would "jump" to question 47 and thus not answering the questions about income and variable costs of the farm.

In the total sample, where no observations is excluded, the mean of the arable land was 64 hectares. The largest farm, with respect to arable land, had 1 250 hectares of arable land. The mean of arable land is 27 hectares when the sample consists of 58 observations. This implies that there are many large farms that have been excluded (with respect to arable land) which could have made a significant difference when estimating the technical efficiency scores. As the larger farms were excluded due to incomplete data, the large farms that did have complete data acted as outliers in the data set. If there had been more large farms, they had not been deemed outliers. However, there would have been a large difference between the farms when 60-70 % of the farms would have a small amount of arable land (less than 20-30 hectares) and about 30-40 % would have a larger amount of arable land (more than 100 hectares). With this in mind, it could have had effects on the estimation of the technical efficiency when there is a cluster of small farms and a cluster of large farms with some farms in between. In the case of the 58 observed farms, a majority of them are relatively small businesses with some farms that are a bit larger.

Another problem with the data was the inconsistent answers from the respondents. In several observations, it was hard to distinguish what type of agricultural production the farm had and if it had any at all, as some respondents would answer that they had agriculture as a main source of income but did not specify what type of production (or the other way around). There was no problem with inconsistent answers when considering questions regarding the horse activities on the farm. 70 % of the sample board horses and where the majority of the respondents have answered what the price of boarding is and how many horses they boarded in 2011. Thus, it would have been easy to calculate the income from horse boarding. However, income from other horse activities as well as variable costs from boarding and other horse activities is not known, and thus many assumptions would be needed in order to isolate the different costs related to boarding and other horse activities.

The Swedish Board of Agriculture (SJV) made large survey of the horse sector in Sweden in 2011. The survey that is used in this study and the one made by SJV is similar in some ways, but there is a difference by the questions asked and who answered the questionnaires. In the survey made by SJV there were no questions related to financials of the horse owner. As there is lacking studies made about the horse sector of Sweden with respect to its economic potential and impact, it would be of interest to conduct another survey where financial questions are included. With consistent economic primary data of the horse owners in Sweden, there could be many studies made to analyze the economic aspects of the Swedish horse sector further.

# 5.2 Econometric results

Even if there are only 58 observations used in order to estimate the technical efficiency of farms that are involved in horse related activities, there are other studies made where the number of observations has been approximately the same or less. In Bravo-Ureta *et al*. (2007) there 6 studies out of 127 using stochastic frontier analysis that have less observations than the number of observations used in this study. One study presented by Bravo-Ureta *et al.* (2007) only has 20 observations. The average mean of technical efficiency for the 6 studies is 77.57, which are larger than the mean technical efficiency score in this study (69.79). The mean technical efficiency scores in the studies analyzed by Bravo-Ureta *et al*. vary between 54 and 89.

It is interesting to note that there is no farm that has a technical efficiency score that is close to the production frontier (in proximity of one). In many studies this is often observed, but not in all studies. The fact that there is no farm that has a technical efficiency score close to 1 does not entirely depend on the functional form used. According to Bravo-Ureta *et al*. (2007), the mean technical efficiency for a Cobb-Douglas functional form is 76.3, which is greater than the mean in this study. They also conclude that one study or more have shown a technical efficiency score of 1. The maximum technical efficiency score assessed by the Cobb-Douglas functional form is the largest compared to those assessed by the translog functional form or other functional forms. However, there is a large variation of technical efficiency scores using the Cobb-Douglas functional form as it also has the lowest technical efficiency score of all functional forms.
The most statistically significant factors of technical inefficiency are $Age^2$, *Age* and *Grant*, both according to the *t*-ratios of the estimates and the estimate of $\gamma$. The interpretation of the value of $\gamma$, the parameter of variance is that it should be close to one (Battese and Coelli, 1995). When $\gamma$ is close to one it indicates that the parameters of technical inefficiency are likely to be highly significant for the analysis of the value of output of the farms. The signs of the factors of $Age^2$ and *Age* implies that younger farmers tend to be more efficient than older ones. This is something that is usually observed in other surveys e.g. Battese and Coelli (1993).

From the second chapter in this study, it can be concluded that there has been increasing returns to scale present in previous studies using stochastic frontier analysis. In one study, increasing returns to scale is statistically significant and in two other studies, increasing returns to scale is present in the production function. In the study made by Bagi *et al*. (1982) it is concluded that small crop farms, related to acreage, exhibit significant increasing returns to scale. Bagi *et al*. (1982) focuses on the whole farm, and could include different types of production activities. This is in line with the results from this study where increasing returns to scale is present in the production function and where farms are relatively small with respect to acreage (the average of the sample farms are below the average of Swedish farms).

The negative value of the estimate of the parameter *Grant* implies that farms that receive any type of grants increase their technical efficiency. The grants are connected to the horse related activities conducted on the farms and the greatest number of grants comes from the EU. The EU-grants for horses are related to the grazing land that the farm has, and are thus applied for at the same time that the farm applies for grants coupled to arable land. It has been concluded in recent studies that farms with lower technical efficiency have had a high dependence on subsidies (e.g. Latruffe, 2010; Latruffe *et al*., 2012; Zhengfei and Lansink, 2006). This is also true for Swedish crop farms (see Zhu and Lansink, 2010). Zhu and Lansink (2010) found further that the share of the crop subsidies in the total subsidies have a positive effect on the technical efficiency of the Swedish farms observed. Zhu and Lansink (2010) assumed that the subsidies were not part of the farm income. In this study, it is also assumed that subsidies are not included in the farm income. The assumption could be made that there are other subsidies present for the farms and thus the subsidies related to horse activities are the share of the total subsidies that have a positive effect on the technical efficiency. However, as the majority of the farms have crop production as their agricultural production it could be the case that subsidies related to crop production could have a positive impact on the technical efficiency of the farms. In order to prove this there is need for further study of the farms by another questionnaire with the appropriate questions for a continued analysis.

# 6 Conclusion

The purpose of this study is to recognize what factors that contributes to a successful farm operation that has multiple outputs. The multiple outputs in this study imply; output from agricultural production and output from horse related activities. The problem formulation is stated below which will be answered in this chapter.

- What is the technical efficiency of farms that have outputs originating from agricultural production and from horse activities?
- What are the factors that contribute to an efficient farm operation that is involved in both agricultural production and horse related activities?

The mean technical efficiency for the farms is 0.69 where the least efficient farm has an efficiency score of 0.51 while the most efficient farm has an efficiency score of 0.83. There is thus no farm that has an efficiency score close to the production frontier, which is the case in many other studies. The efficiency scores are all clustered together around the mean. It could thus be concluded that there is no farm that is in close proximity to be inefficient, but there is on the other hand no farm that is fully efficient either.

The factors that contribute to an efficient farm operation are derived from the determinants of technical inefficiency. Factors that do contribute to an efficient farm operation that is involved in horse related activities are farmers that receive grants. The factor age does also contribute to technical efficiency where younger farmers tend to be more efficient than older farmers do.

Due to the results in this study, there is one question that arise and which could be answered through further study and analysis of farms that are involved in horse related activities. The question is regarding the subsidies: are there any other subsidies present that are related to the agricultural production on the farms? With the positive effect the subsidies related to horses have on technical efficiency it would be of interest to estimate the effect a subsidy related to agricultural production have on the technical efficiency of the farms.

# Bibliography

## Literature and publications


Ahmad, M. and Bravo-Ureta, B. (1996). Technical efficiency measures for dairy farms using panel data: a comparison of alternative model specifications. *Journal of Productivity Analysis*, 7: 399-416.

Aigner, D., Lovell, C. A. K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production functions. *Journal of Econometrics*, 6: 21-37.

Bagi, F. (1982). Relationship between farm size and technical efficiency in West Tennessee Agriculture. *Southern Journal of Agricultural Economics*, 14: 139-144.

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: 153-169.

Battese, G. E. and Tessema, G. (1993). Estimation of stochastic frontier production functions with time-varying parameters and technical efficiencies using panel data from Indian villages. *Agricultural Economics*, 9: 313-333.

Battese, G. E. and Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20: 325-332.

Battese, G. E. and Broca, S. S. (1997). Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. *Journal of Productivity Analysis*, 8: 395-414.

Battese, G. E. (1992). Frontier production functions and technical efficiency: a survey of empirical applications in agricultural economics. *Agricultural Economics*, 7: 185-208.

Battese, G. E. (1997). A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values. *Journal of Agricultural Economics*, 48(2): 250-252.

Binam, J., Tonye, J., Njankoua, W., Nyambi, G. and Akoa, M. (2004). Factors affecting the technical efficiency among smallholder farmers in the slash and burn agriculture zone of Cameroon. *Food Policy*, 29: 531-545.

Bogetoft, P. and Otto, L. (2011). *Benchmarking with DEA, SFA, and R.* New York: Springer Science+Business Media.

Bravo-Ureta, E. B., Solís, D., López, M. H. V., Maripani, F. J., Thiam, A. and Rivas, T. (2007). Technical efficiency in farming: a meta-regression analysis. *Journal of Productivity Analysis*, 27(1): 57-72.

Bravo-Ureta, E. B. and Pinheiro, A. E. (1993). Efficiency analysis of developing country agriculture: a review of the frontier function literature. *Agricultural and Resource Economics Review*, 22: 88-101.

Bravo-Ureta E. B. and Pinheiro, A. (1997). Technical, economic, and allocative efficiency in peasant farming: evidence from the Dominican Republic. *Development Economics*, 35: 48-67.

Brümmer B. and Loy, J. (2000). The technical efficiency impact of farm credit programs: a case study of Northern Germany. *Journal of Agricultural Economics*, 51: 405-418.

Brümmer B. (2001). Estimating confidence intervals for technical efficiency: the case of private farms in Slovenia. *European Review of Agricultural Economics*, 28: 285-306.

Coelli, T. J. (1996). A Guide to FRONTIER version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation. CEPA Working Papers No. 7/96, Armidale: University of New England.

Coelli, T. and Perelman, S. (1996). Efficiency measurement, multiple-output technologies and distance functions: with application to European railways. CREPP Discussion Paper No. 19/05, Liége: Université de Liége.

Coelli, T. and Fleming, E. (2004). Diversification economies and specialisation efficiencies in a mixed food and coffee smallholder farming system in Papua New Guinea. *Agricultural Economics*, 31: 229-239.

Dumler, J. T., Burton, O. R. Jr. and Kastens, L. T. (2003). Predicting farm tractor values through alternative depreciation methods. *Review of Agricultural Economics*, 25(2): 506-522.

Emvalomatis, G., Lansink, A. O. and Stefanou, S. E. (2008). An examination of the relationship between subsidies on production and technical efficiency in agriculture: The case of cotton producers in Greece. Paper prepared for presentation at the 107th EAAE Seminar "*Modelling and Rural Development Policies*", Spain: Sevilla.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3): 253-290.

Fox, J. (1997). *Applied regression analysis, linear models, and related methods.* California: Sage publications.

Furugren, B. (1990). *Hästens historia (History of the horse)*. Uppsala: Sveriges lantbruksuniversitet, Hippologiska högskoleutbildningen.

Giannakas, K., Tran, K. and Tzouvelekas, V. (2000). Efficiency, technological change and output growth in Greek olive growing farms: a box-cox approach. *Applied Economics*, 32: 909-916.

Greene, W. H. (1990). A gamma distributed stochastic frontier model. *Journal of econometrics*, 46: 141-164.

Gujarati, D. N. and Porter, D. C. (2009). *Basic econometrics (fifth edition)*. New York: McGraw-Hill/Irwin.

Hadri, K., Guermat, C. and Whittaker, J. (2003). Estimating farm efficiency in the presence of double heteroscedasticity using panel data. *Journal of Applied Economics*, 6: 255-268.

Hadri, K., Guermat, C. and Whittaker, J. (2003). Estimation of technical inefficiency effects using panel data and doubly heteroscedastic stochastic production frontiers. *Empirical Economics*, 28: 203-222.

Hedberg, A. (2002). *Hästens roll för samhälle och lantbruk* (*The horse's part for the society and agriculture*), Uppsala: Sveriges lantbruksuniversitet, Institutionen för ekonomi.

Heshmati, A. and Kumbhakar, S. C. (1997). Estimation of technical efficiency in Swedish crop farms: a pseudo panel data approach. *Journal of Agricultural Economics*, 28: 22-37.

Huang, C. and Bagi, F. (1984). Technical efficiency on individual farms in Northwest India. *Southern Economic Journal*, 50: 108-115.

Huang, C. J. and Liu, J. T. (1994). Estimation of a non-neutral stochastic frontier production function. *Journal of Productivity Analysis*, 5: 171-180.

Ivaldi, M., Perrigne, I. and Simioni, M. (1994). Productive efficiency of French grain producers: a latent variable model. *Journal of Productivity Analysis*, 5: 287-299.

Johansson, D., Andersson, H. and Hedberg, A. (2004). *Hästnäringens samhällsekonomiska betydelse i Sverige (The economic importance of the horse sector in Sweden)*, Uppsala: Repro, Sveriges lantbruksuniversitet.

Jordbruksverket. (2012). *Hästhållning i Sverige 2010* (*Horse owners in Sweden 2010*), rapport 2012:1. Jönköping: Jordbruksverket.

Jordbruksverket. (2011). Statistiskt meddelande JO 24 SM 1101 *Antalet hästar och anläggningar med häst 2010* (Statistical message JO 24 SM 1101 *Horses and horse establishments in 2010*), published 2011-01-20.

Khumbakar, S. C. and Lovel, C. A. K. (2000). *Stochastic Frontier Analysis*, Cambridge: The press syndicate of the University of Cambridge.

Koop, R. J. and Smith V. K. (1980). Frontier production function estimates for steam electric generation: a comparative analysis. *Southern Economic Journal*, 47: 1049-1059.

Kurkalova, L. and Carriquiry, A. (2003). Input- and output-oriented technical efficiency of Ukrainian collective farms, 1989-1992: Bayesian analysis of stochastic production frontier model. *Journal of Productivity Analysis*, 20: 191-211.

Lansink, A. (2000). Productivity growth and efficiency measurement: a dual approach. *European Review of Agricultural Economics*, 27: 59-73.

Latruffe, L. (2010). Competitiveness, productivity and efficiency in the agricultural and the agri-food sectors. OECD Food, Agriculture and Fisheries Working Paper: OECD Publishing.

Latruffe, L., Bravo-Ureta, B. E., Moreira, V. H. Desjeux, Y. and Dupraz, P. (2012). Productivity and subsidies in the European Union: An analysis for dairy farms using input distance frontiers. *International Association of Agricultural Economists Triennial Conference*. Foz do Iguacu, Brazil.

Lantmannen redskap. (2005). Stockholm: LRF Media AB.

Liljenstolpe, C. (2009). *Horses in Europe*. Report presented at EU Equus conference 2009.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18: 435-444.

Myrdal, J. (2001). Från adelsattribut till tonårsvurm. I: *Människan och naturen. Etnobiologi i Sverige* (From an aristocracy accessory to a teenage obsession. In: *Mankind and nature. Ethnobiology in Sweden*), volym 1:315-321. Stockholm: Wahlström and Widstrand.

OECD. (2001). *Multifunctionality Towards an Analytical Framework*, Paris: OECD.

Rezitis, N. A., Tsiboukas, K. and Tsoukalas, S. (2002). Measuring technical efficiency in the Greek agricultural sector. *Applied Economics*, 34: 1345-1357.

Rezitis, N. A., Tsiboukas, K. and Tsoukalas, S. (2003). Investigation of factors influencing the technical efficiency of agricultural producers participating in farm credit programs: the case of Greece. *Journal of Agricultural Applied Economics*, 35: 529-541.

Schmidt, P. (1977). *Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers*. Workshop paper number 7702, Department of Economics, Michigan State University.

Skatteverket. (2010). *Handledning för fastighetstaxering 2011 (Introduction to assessment of taxation value)*, Västerås: Edita.

SLU. (2001). *The horse industry in the European Union.* Final report from EU Equus 2001. Uppsala: Sveriges lantbruksuniversitet, Institutionen för ekonomi.

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of econometrics*, 13: 57-66.

Tell, M. (2001). Lantbrukets guldkalv. *Lantbruk*, nr 37/2001 (Bilaga häst): 3.

Thiam, A., Bravo-Ureta, E. B. and Rivas, T. E. (2001). Technical efficiency in developing country agriculture: a meta-analysis. *Agricultural Economics*, 25: 235-243.

Tzouvelekas, V., Pantzios, J. C. and Fotopoulos, C. (2001). Technical efficiency of alternative farming systems: the case of Greek organic and conventional Olive-growing farms. *Food Policy*, 29: 546-569.

Tzouvelekas, V., Pantzios, J. C. and Fotopoulos, C. (2001). Economic efficiency in organic farming: evidence from cotton farms in Viotia, Greece. *Journal of Agricultural Applied Economics*, 33: 35-48.

Zhengfei, G. and Lansink, A. O. (2006). The source of productivity growth in Dutch agriculture: A perspective from finance. *American Journal of Agricultural Economics*, 88: 644-656.

Zhu, X. and Lansink A. O. (2010). Impact of CAP subsidies on technical efficiency of crop farms in Germany, the Netherlands and Sweden. *Journal of Agricultural Economics*, 61: 545-564.

## Internet

Agriwise, Verktyg för ekonomisk planering och analys, www.agriwise.org
1. *Databoken 2013 (Database 2013)*, 2013-03-07

EC, European Commission, ec.europea.eu
1. FADN - Public database - Farm economy focus - Sweden, 2014-08-26
http://ec.europa.eu/agriculture/rica/database/factsheets_en.cfm

HNS, Swedish Horse Foundation, www.nshorse.se
1. *Forskningsanslag (Research grants)*, 2013-02-13
http://www.nshorse.se/polopoly_fs/1.57354.1356036092!/beslutslista_shf_121218.pdf

Skatteverket, The Swedish Enforcement Administration, www.skatteverket.se
1. *Omräkningskruser för euro per redovisningsperiod*, 2014-08-26
https://www.skatteverket.se/foretagorganisationer/foretagare/euronskatterna/omrakningskurser/redovisningsperioder.4.2ef18e6a125660db8b080004155.html

SCB, Statistics Sweden, www.scb.se
1. *Statistical database*, 2013-02-12
http://www.scb.se/Pages/SSD/SSD_TablePresentation____340486.aspx?rxid=3fe9799a-ff8a-40c1-963f-ab98f79b0168andlayout=tableViewLayout1

SJV, Swedish Board of Agriculture, www.jordbruksverket.se
1. *Statistics*, 2013-02-13
http://statistik.sjv.se/Database/Jordbruksverket/Jordbruksforetag/Jordbruksforetag.asp

2. *Statistics*, 2013-02-13
http://statistik.sjv.se/Database/Jordbruksverket/Husdjur/Husdjur.asp

TSP, TSP international, www.tspintl.com
1. *Info - What is TSP?*, 2013-05-13
http://www.tspintl.com/tspinfo/whatis.htm

# Appendix A: Questionnaire on horse boarding

**You are**
☐Man ☐Woman

**2. Age _____________________**

**3. Your role in the business:**____________________________________________________________

**4. What education level do you have?**
☐ Primarily school ☐ Secondary school ☐ Post-secondary education (Not University)
☐ University

**5. Do you have any horse related education?** (Check one or more options)

☐ Agricultural secondary school
☐ Equine education
☐ Veterinary education
☐ Agronomist
☐ Feed knowledge (courses, etc.)
☐ Referee education
☐ Blacksmith
☐ I have no horse related education
☐ Horse trainer education
☐ Riding instructor education
☐ Vocational qualification
☐ Entrepreneurship
☐ Business administration
☐ Breeding related education
☐ Groom education
☐ Other, _____________________

**6. Where is your farm located?**
☐ In a rural area ☐ In a peri-urban area ☐ In an urban area

**7. How far is it to the closest city/town with at least 5000 inhabitants? ______km**

**8. How many individuals live in this city/town?__________________individuals**

**9. What year did you buy the farm?** _____________________________________________

**10. What year did you start your business in the horse sector?** ________________________________

11. **What kind of business entity is the riding school?**

☐ Limited company
☐ General partnership
☐ Sole proprietorship
☐ Other

**12.What production activities are conducted on the farm?**
☐Cereals
☐ Gardening
☐ Beef
☐ Contract-work
☐ Forestry
☐ Dairy
☐ Poultry
☐ Pork
☐ Estate management
☐ Other services

☐ Horse activities ☐Other, namely___________________________

**13. How many full-time positions do you have at the riding school?** (Including yourself)_______

**14. How many people share these positions?** _____________________

**15. What kind of machinery do you have in your business?**

| Machine | Horse power | Year of construction | width |
|---|---|---|---|
| **Tractor 1** | | | |
| **Tractor 2** | | | |
| **Tractor 3** | | | |
| **Tractor 4** | | | |
| **Loader** | | | |
| **Treshing machine 1** | | | |
| **Treshing machine 2** | | | |
| **Seed drill** | | | |
| **Harrow** | | | |
| **Cultivator** | | | |
| **Plough** | | | |
| **Peast spray** | | | |
| **Strawpress** | | | |

**16.What type of buildings do the farm have?**

| Type of buildings | Size of buildings |
|---|---|
| ☐ Farm building for beef, poultry or pork production | Kvm |
| ☐ Stable | Kvm |
| ☐ Riding hall | Kvm |
| ☐ Barn | Kvm |
| ☐ Machine hall | Kvm |
| ☐ Workshop | Kvm |
| ☐ Storage for cereals | M3 |
| ☐ Storage for vegetables | M3 |

Notes: Kvm = square meters and M3 = cubic meters

**17.How many hectares do the farm operate?**

Arable land ___________ ha of which is rented__________ ha

Grazing land ___________ ha of which is rented__________ ha

Forest ___________ ha

**Information about your horse business**

**18. What year did you begin your establishment in the horse sector?____________**

**19. What type of horse activities are included in your enterprise?**

☐ Training of trotting horses ☐ Trail riding
☐ Breeding ☐ Horse boarding
☐ Riding school ☐ Tourism
☐ Recreational purposes ☐ Agriculture
☐ Goods and services for the horse sector

**20. Is your breeding more hobby or business oriented?**
□ Hobby
□ Business
□ Other__________________________________________________

**21. Is the horse business your main source of income?**
□ Yes
□ NO, my main source of income is: ________________________________

**22. How large a part of your income comes from the horse activities?**

☐ 0-10% ☐ 50-60%
☐ 10-20% ☐ 60-70%
☐ 20-30% ☐ 70-80%
☐ 30-40% ☐ 80-90%
☐ 40-50% ☐ 90-100%

**23. How many hours in total do you spend on your horse related activities? _________hours**

**24. Have you applied for any grants for your horse business?**

☐ No ☐ Idrottslyftet
☐ Business grant ☐ LOK-stöd
☐ Project grant ☐ Community grant
☐ Leader ☐ EU-grant
☐ Environmental grant ☐ Other____________________

**25. What type of horse breed is found on your farm? The number of each breed.**

☐ Warm-blooded ________________st ☐ Shetlandic pony ________________st
☐ Coldblooded ________________st ☐ Icelandic __________________st ☐
Other pony ________________st ☐ Other ____________________st

**26. What is the capacity of horses in your stable? ____________st**

**27. What was the average number of horses on your farm in 2011?** _________st.
**… of these where**
_________ own
_________ boarded horses
_________ grazing
_________ other, namely: ________________________________________

**28. What type of layout do you have for the horses on the farm?**

☐ Stable with boxes: Nr of horses_________________
☐ Stable with boxes: Nr of horses_________________
☐ Loose Nr of horses_________________

**29. What is the market value of your own horses? _______________** SEK

**30. What is the interest of boarding horses in your area?**
☐ Very low ☐ High
☐ Low ☐ Very high
☐Do not know

**31. What other facilities are in boundary to your farm?**

☐ Riding hall
☐ Trails
☐ Pens
☐ Summer grazing
☐ Walk-machine
☐ Trotting-track
☐ Track with obstacles
☐ Other, __________

**32. Do you board horses?** ☐ Yes ☐ No **(If your answer is no, please proceed to question 47)**

**33. If you board horses, what is the price for the following services?**

| | SEK per month per horse | SEK per month per ponny |
|---|---|---|
| **Box** | | |
| **Box with window** | | |
| **Box with paddock** | | |
| **Grazing** | | |
| **Other, specify** | | |

**34. Can the owner of the horse decide about the level of services?**

☐ No
☐ Yes, by contract
☐ Yes, orally

**35. If the horse owner cannot decide about the level of services, what is then included in the rent?**

☐ Feed
☐ Summer grazing
☐ Inlet and discharge
☐ Pens
☐ Riding hall
☐ Solarium
☐ De-worming
☐ Mucking out
☐ Litter
☐ On and off of blankets
☐ Training
☐ Vetrenary
☐ Mucking box
☐ Other_____________________________

**36. If the horse owner is able to decide what is included in the rent, what does the owner want to include in the rent?**

| **Attribute** | **SEK per month** | **Attribute** | **SEK per month** |
|---|---|---|---|
| Feeding | | Litter | |
| Mucking | | Inlet and discharge | |
| Riding hall | | On and off of blankets | |
| Training | | Pens | |
| Veterinary | | Summer grazing | |
| Solarium | | De-worming | |
| Mucking box | | Other_____________ | |

**37. How do you set the price level for boarding?**

☐ By comparing price level with the competition of others
☐ By negotiation with the horse owner
☐ By calculating the costs of boarding horses

**38. How many facilities boarding horses that are in the vicinity of your farm (10 kilometers)?**

☐ Fewer than 5
☐ 5-10
☐ 10-15
☐ 15-20
☐ 20-25
☐ No one

**39. What was the income from your farm and horse activities?**

| Category | Income (SEK) |
|---|---|
| Total income | |
| Total income only from horse activities | |

**40. What were the variable costs of you farm and horse activities?**

| Category | Variable cost (SEK) |
|---|---|
| Total variable costs | |
| Total variable costs for the horse activities | |

**41. What was the cost of labor?**

| Category | Cost (SEK) |
|---|---|
| Total cost of labor | |
| Cost of labor related to horse activities | |

**42. What do you find the quality of your facilities to be?**

| | Very good | Good | Bad | Very bad |
|---|---|---|---|---|
| **Space for boarding** | | | | |
| **Quality of training facility** | | | | |
| **Possibility for grazing** | | | | |
| **Number of trails** | | | | |
| **Parking options** | | | | |
| **Parking options for horse coach** | | | | |

**43. How do you reach your customers?**

☐ Signs
☐ Homepage
☐ Own logo
☐ Other, namely_______________________
☐ Folders
☐ Advertisement
☐ Cooperation with riding schools

**44. Compared to other businesses with horse boarding, how much do you advertise?**

☐ A lot more ☐ More ☐ Equally ☐ Less ☐ Much less

**45. What type of customers do you have?**

☐ Professional riders ☐ Hobby riders

**46. In what area does your business have its strengths and weaknesses?**

| Category | Very good | Good | Less good | Bad |
|---|---|---|---|---|
| **Price of boarding** | | | | |
| **Level of services** | | | | |
| **Riding possibilities** | | | | |
| **Meeting the customers whishes** | | | | |
| **Usage of Internet** | | | | |
| **Usage of advertisement** | | | | |

**47. Please answer the following statements:**

| | *Does not acknowledge* | *Acknowledge partly* | *Acknowledge* | *Acknowledge greatly* | *Acknowledge completely* |
|---|---|---|---|---|---|
| I like to have a lot of people close to me | ☐ | ☐ | ☐ | ☐ | ☐ |
| I like routines | ☐ | ☐ | ☐ | ☐ | ☐ |
| The work in the horse sector is important to me | ☐ | ☐ | ☐ | ☐ | ☐ |
| The income from the different activities are important to me | ☐ | ☐ | ☐ | ☐ | ☐ |
| The development of the business is important to me | ☐ | ☐ | ☐ | ☐ | ☐ |

**48. What is the level of significance for your business to receive information from these different sources?**

| **Source of information** | *No significance* | *Less significant* | *Significant* | *Greatly significant* | *Highly significant* |
|---|---|---|---|---|---|
| Media | ☐ | ☐ | ☐ | ☐ | ☐ |
| Internet | ☐ | ☐ | ☐ | ☐ | ☐ |
| Employees | ☐ | ☐ | ☐ | ☐ | ☐ |
| Family | ☐ | ☐ | ☐ | ☐ | ☐ |
| Other riding schools | ☐ | ☐ | ☐ | ☐ | ☐ |
| Own experience | ☐ | ☐ | ☐ | ☐ | ☐ |
| Riksidrottsförbundet[24] | ☐ | ☐ | ☐ | ☐ | ☐ |
| SvRF[25] | ☐ | ☐ | ☐ | ☐ | ☐ |
| Jordbruksverket | ☐ | ☐ | ☐ | ☐ | ☐ |
| LRF | ☐ | ☐ | ☐ | ☐ | ☐ |
| The municipality | ☐ | ☐ | ☐ | ☐ | ☐ |
| Other | ☐ | ☐ | ☐ | ☐ | ☐ |

**49. Do you have economic goals for your horse establishment?**

☐ Yes, on paper
☐ Yes, in my head
☐ No

| | *Very well* | *Well* | *Satisfactory* | *Less well* | *Badly* |
|---|---|---|---|---|---|
| **50. If yes, do you reach your goals?** | ☐ | ☐ | ☐ | ☐ | ☐ |

**51. Do you see any troubles regarding development of your horse establishment?**

☐ No

☐ Yes, namely__________________________________________________

[24] The Swedish Sports Confederation
[25] The Swedish Equestrian Federation

**52. How does your horse business look like in 5 year's time?**

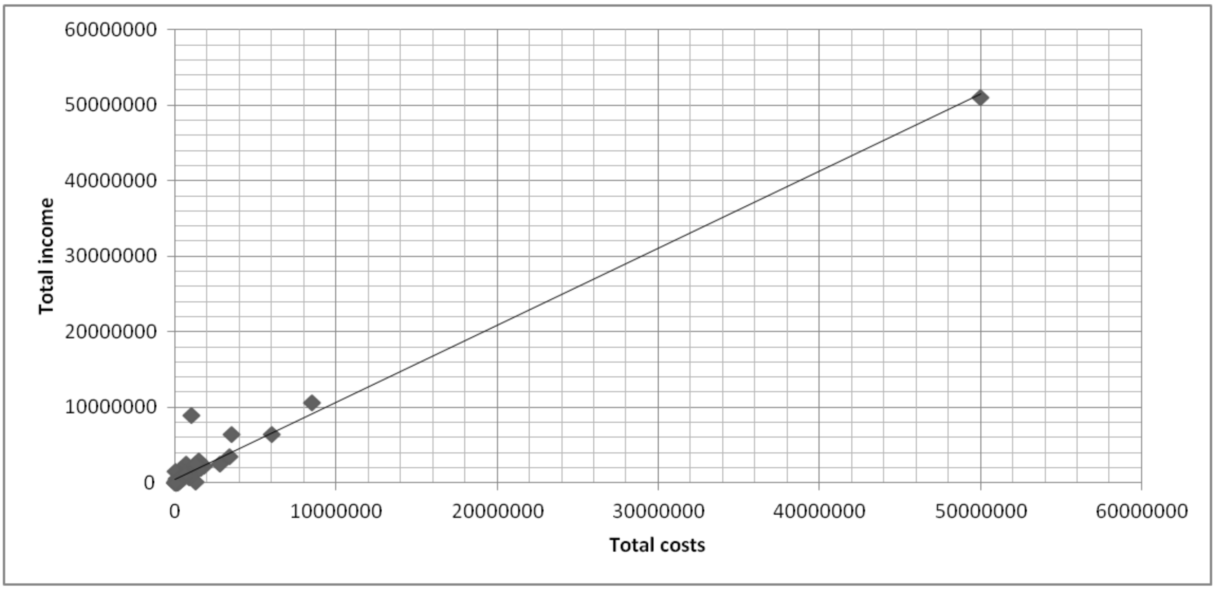
☐ Increased significantly
☐ Increased somewhat
☐ As now
☐ Decreased somewhat
☐Ended the establishment the business is terminated.

**53. How do you think your level of debt is?**

☐ Too high ☐ all right

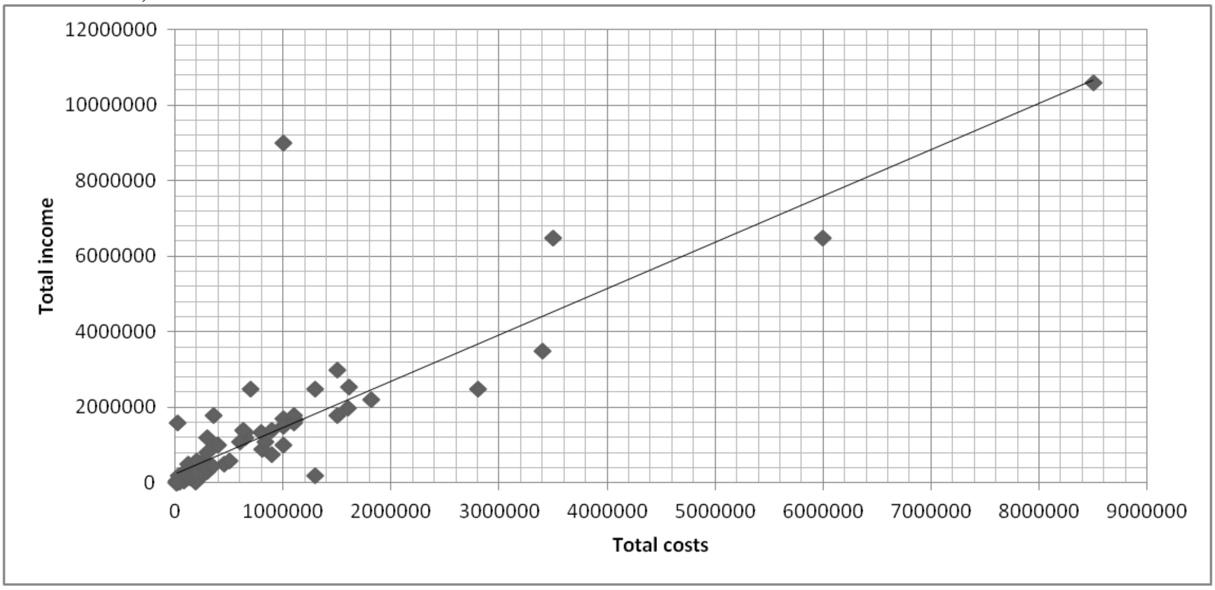
# Appendix B: Figures showing evident outliers in the data

*Figure B.1. Total income vs. total costs (N=113). (Source: Boarding data with own calculation).*




In the figure B.1 there are two observations that have an income greater or equal to 50 million SEK. These two observations are excluded from the dataset.

*Figure B.2. Total income vs. total costs (N=111). (Source: Boarding data with own calculation).*




In figure B.2. there are still outliers present as well as observations with incomplete data. In the end, there will be observations that have an income less than 2 million SEK and costs less than 1.8 million SEK.

# Appendix C: Remaining percentage value

The depreciation method is used in order to find how much the value of a tractor has decreased depending on how old the tractor is (Dumler *et al.*, 2003). The most common and most used method is the remaining value percentage, also referred to *RVP*. The formula of the remaining value percentage equals:

$$\mathrm{RVP} = 0{,}68(0{,}92)^{\mathrm{n}}\ , \tag{C.1}$$

where *n* equals the age of the machinery in years.

# Appendix D: Assessment of taxation value

The assessment of the taxation value of a building is done by the following formula:

$$R = N \cdot F_e \,, \qquad \text{(D.1)}$$

where $R$ equals the value of the building, $N$ equals the E-factor for the area the building is located and $F_e$ equals the relative value according to a E-table that states the size, age and nature of the building (Skatteverket, 2010).

The E-tables for different types of agricultural buildings are found in the appendix of the book by Skatteverket. The E-factors are found on Skatteverkets homepage, where the values are listed after each municipality in Sweden.