Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

BACHELOR THESIS

Determinants of Technical Efficiency and its Change among Ghanaian Cocoa Farmers

Author: **David Valenta**

Supervisor: **Petr Janský Ph.D.**

Academic Year: **2013/2014**

Declaration of Authorship

1. Hereby I declare that I have compiled this bachelor thesis independently, using only the listed literature and sources.
2. I declare that the thesis has not been used for obtaining another title.
3. I agree on making this thesis accessible for study and research purposes.

Prague, 16th May 2014

____
Signature

Acknowledgments

I am grateful to Petr Janský Ph.D. for his patient supervision and useful comments and to everyone else who helped me and supported me while I was writing this thesis. My special thanks go to the Centre for the Study of African Economies, Oxford University for providing me with the data.

Abstract

This thesis analyses determinants of technical efficiency (TE), its change and scale efficiency on an unbalanced panel of more than 700 Ghanaian cocoa farmers collected between 2004 and 2010. Using variable returns to scale data envelopment analysis the TE scores are obtained and the mean TE found is 35.4%. In second stage, random effects Tobit regression is used to investigate the determinants of efficiency and number of significant determinants is found from categories of farmers' characteristics, different technical specifications of the farm, focus on farming, farm management style and investments. Moreover, this thesis contributes to the efficiency debate by researching determinants of efficiency change, finding gender and savings to have significant positive effect among others.

JEL Classification	D24, O13, Q12
Keywords	DEA, Farmers' Efficiency, Efficiency Change, Determinants of Efficiency, Ghana, Cocoa

Author's e-mail	valentadavid@gmail.com
Supervisor's e-mail	jansky@fsv.cuni.cz

Abstrakt

Tato práce analyzuje determinanty technické efektivity (TE), její změny a efektivity rozsahu na nevyrovnaném panelu více než 700 ghanských farmářů, který byl posbírán mezi lety 2004 a 2010. Hodnoty TE byly nalezeny pomocí "data envelopement analysis" modelu s proměnlivými výnosy z rozsahu. Průměrná hodnota TE nalezena činila 35.4%. V druhé fázi byly vyzkoumány determinanty efektivity pomocí Tobitovy regrese s náhodnými efekty a bylo nalezeno mnoho proměnných se signifikantním efektem v kategoriích jako: charakteristiky farmáře, technické specifikace farmy, zaměření na farmaření, styl managementu a investic. Navíc, tato práce přispívá do debaty o efektivitě tím, že zkoumá determinanty změny efektivity a nalezla, mimo jiné, že pohlaví farmáře a spoření mají signifikantní pozitivní vliv na změnu efektivity v čase.

Klasifikace JEL	D24, O13, Q12
Klíčová slova	DEA, efektivita farmářů, změna efektivity, determinanty efektivity, Ghana, kakao

E-mail autora	valentadavid@gmail.com
E-mail vedoucího práce	jansky@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

COCOBOD Ghana Cocoa Board
CRS Constant returns to scale
CSAE Centre for the Study of African Economies
DEA Data envelopement analysis
DMU Decision making unit
GCFS Ghana Cocoa Farmers Survay
GDP Gross domestic product
GLSS Ghana Living Standars Survay
GSS Ghana Statistical Service
SE Scale efficiency
SFA Stochastic frontier analysis
TE Technical efficiency
VRS Variable returns to scale

Bachelor Thesis Proposal

Author	David Valenta
Supervisor	Petr Janský Ph.D.
Proposed topic	Efficiency Analysis of Ghana Cocoa Farmers

Core bibliography

1. BINAM, J. N., Kalilou SYLLA, Ibrahim DIARRA a Gwendoline NYAMBI. Factors Affecting Technical Efficiency among Coffee Farmers in Côte dIvoire: Evidence from the Centre West Region. African Development Review. 2003, roč. 15, č. 1.

2. BINAM, J. N., Jim GOCKOWSKI a Guy Blaise NKAMLEU. Technical efficiency and productivity potential of cocoa farmers in West African countries. The Developing Economies. 2008, roč. 46, č. 3.

3. BRAVO-URETA, Boris E. a Robert E. EVENSON. Efficiency in agriculturalproduction: The case of peasant farmers in eastern Paraguay. Agricultural Economics. 1994, ro. 10, . 1. BRAVO-URETA, Boris E., et. al., Technical efficiency in farming: a meta-regression analysis. Journal of Productivity Analysis. 2007, roč. 27, č. 1.

4. COELLI, Tim, Sanzidur RAHMAN a Colin THIRTLE. Technical, Allocative, Cost and Scale Efficiencies in Bangladesh Rice Cultivation: A Non-parametric Approach. Journal of Agricultural Economics. 2002, roč. 53, č. 3.

5. COELLI, T.J, RAO, D.S.P., O'DONNELL, C.J. a BATTESE, G.E. An Introduction to Efficiency and Productivity Analysis. 2. vyd. Springer, 2005. ISBN 978-0-387-24265-1.

6. RICHMAN, Dzene. What drives efficiency on the Ghanian cocoa farm?. In: CSAE Conference 2010 Economic Development in Africa [online]. 2010.

Dostupné z:
http://www.csae.ox.ac.uk/conferences/2010-EDiA/papers/498-Dzene.pdf

7. THIAM, Abdourahmane, Boris E. BRAVO-URETA a Teodoro E. RIVAS. Technical efficiency in developing country agriculture: a meta-analysis. Agricultural Economics. 2001, roč. 25, 2-3.

Preliminary scope of work Productivity growth in agriculture in coming years will be important for development in developing countries. Productivity growth can be divided into technology change and efficiency growth. In my thesis I would like to examine technical efficiency (TE) of Ghana cocoa farmers. Previous studies implies, that theirs TE is not high, Binam, Grockowski and Nkanleu (2008) in their study determined TE to be 44%, that is even less than their results from other West African countries. And also less than usual efficiency of farmers in Africa and low-income countries in general, which in their study interested in articles determining productivities in agriculture was found by Bravo-Ureta et. al. (2007) to be 73,7% and 74,1% respectively. So there is room for improvement.

On rich dataset collected by CSAE from hundreds of farmers, I will determine efficiencies of particular farmers using data envelopment analysis. Then I will examine influences of various technical and socio-economic factors on efficiency and its change in time, similarly as Coelli et. al. (2002) on Bangladesh rice farmers and Binam et. al (2003) on Côte d'Ivoire coffee farmers using DEA or Richman (2010), who studied efficiency of Ghana cocoa farmers and Bravo-Ureta, Everson (1994), researching efficiency of farmers in Paraguay, using stochastic frontier.

In my thesis I want to find, how to increase efficiency of these farmers and current results compare with existing literature interested in this topic, e.g. Richman (2010). Beyond that I would like to examine what influences efficiency change of particular farmers in time. Possible answers to how increase efficiency could be increased education, better availability of financial services, change in the way the farm is managed (e.g. the way the employees are paid) and other.

Chapter 1
Introduction

Agriculture and cocoa production play important part in Ghanaian economy. Ghana is the second largest cocoa producer in the world (International Cocoa Organization (2014)). It is suggested that Ghanaian cocoa sector is only having lower yields and levels of efficiency then it is in the neighbouring countries.

The panel used in this thesis was made as part of the Ghana Cocoa Farmers Survey (GCFS). The sample was collected bi-annually between years 2004 and 2010 by researches from the Centre for the Study of African Economies (CSAE), the Ghana Cocoa Board (COCOBOD) and the Ghana Statistical Service (GSS). The sample consist of more than 700 farmers and total of about 1500 observations in several Ghanaian regions.

Using two stage method this thesis will obtain technical efficiency scores using variable returns to scale data envelopment analysis in the first phase. In the second stage then we will research the determinants of efficiency using random effects Tobit regression. Next, this thesis will research previously neglected topic of efficiency change in time and present some completely new findings, by investigating the determinants of efficiency change. Last, scale efficiency will also be discussed.

Among the possible determinants of efficiency that this thesis will research whether they have some significant effects will be various farmer characteristics, different technical specifications of the farm, focus on farming, farm management style and investments. The findings of this thesis will then be compared with findings of previous studies on the efficiency of Ghanaian cocoa farmers, like Amos (2007), Danso-Abbeam et al. (2012), Onumah et al. (2013) and Richman (2010).

Chapter 2 of this thesis will give reader some background on the problem; Chapter 3 will present the methodology and Chapter 4 the data. The results will be presented in Chapter 5 and last chapter will conclude the findings.

Chapter 2
Background of the Study

2.1. Cocoa Production in Ghana

Most of the worlds total cocoa production, above 70%, is produced in the West Africa. Ghana is the country with second-highest cocoa production in the world, with only neighbouring Côte d'Ivoire producing more (International Cocoa Organization (2014)).

In Ghana, same as in many other African and other developing countries, agriculture plays much more important role in the economy then in the developed countries. According to World Bank (2014), agriculture made about a quarter of the Ghanaian GDP in 2012, almost as much as industry did. The share of agriculture in Ghanaian economy is steadily declining from 32% in 2009 to the 23% in 2012 with growth averaging just about 4% between 2006 and 2009. On the other hand, the proportion of population employed in agriculture remains high, being 42% in 2010 (World Bank (2013a)).

Cocoa production plays important role in Ghanaian agriculture, but with its share steadily decreasing from just below 13% in 2005 to 11.5% in 2009 (Ghana Statistical Service (2014)). However, the total cocoa production in Ghana has still risen in the last decade (Ghana Cocoa Board (2014)).

Still, it seems the cocoa sector in Ghana is chronically under-achieving its potential. While the experimental study from 1970's suggests potential yield of about 2500 kg per ha a year (refer for example to Teal and Vigneri (2004)) the studies of Ghanaian cocoa production like Zeitlin (2005) and Richman (2010) on dataset from same survey as this thesis will use, or Danso-Abbeam et al. (2012) and Aneani et al. (2011) on different datasets, suggest the mean yields far below, less than 20%, of the experimental amount. And these yields , not being even 400 kg per ha, are also much lower then the ones observed in other countries, like about 590 kg per ha observed by Eyitayo et al. (2011) in Nigeria. This suggests big room for improvement in Ghanaian cocoa production that could have significant positive impact on lives of many Ghanaians.

2.2. Efficiency Studies

There are many studies on efficiency of farmers. Majority uses stochastic frontier analysis (SFA) to analyse efficiency, however, many use data envelopment analysis (DEA). Some examples of studies using SFA are for example Amos (2007) or Danso-Abbeam et al. (2012), and some using DEA are Helfand and Levine (2004) or Coelli et al. (2002).

Some studies researching cocoa farmers efficiency in Ghana are for example Danso-Abbeam et al. (2012) and Richman (2010), they found mean technical efficiencies of 49% and 44% respectively. These results are somewhat lower than the ones found for cocoa farmers in Nigeria, where Amos (2007) and Eyitayo et al. (2011) found mean TE of 72% and 87% respectively. Moreover, Binam et al. (2008) in his cross-county comparison founds Ghanaian cocoa farmers to be the least efficient ones in West Africa. Ghanaian farmers were found to have TE of only 44%, while farmers in Côte d'Ivoire, Cameroon and Nigeria had 58%, 65% and 74% respectively. Moreover, for example Bravo-Ureta et al. (2007) in his work looking at other efficiency studies found that mean TE found by studies concerning African agriculture is around 74%. This suggests Ghanaian cocoa sector is doing worse then the ones in neighbouring countries.

Number of studies also looks for the determinants of efficiency. Among the farmers' characteristics being mentioned to have an effect on technical efficiency are for example: the age of farmer, gender, farmers' education or experience. Some other factors influencing efficiency mentioned are for example: the size of the farm, access to credit, negative events on the farm, different kinds of technology or for example seed variety. Some of these effects will be more in detail mentioned later and compared with our results.

Some authors, like Onumah et al. (2013) choose to include some of these variables, especially the technical ones, directly into the frontier estimation and only look for the mean TE scores and social determinants. However, many, same as we will, choose to only investigate these effects once the TE scores are obtained. This approach was used for example by Richman (2010) or Coelli et al. (2002). This results are easier to interpret and consider some negative effect for example of the age of trees as an inefficiency, which it is, and not as part of production frontier.

2.3. Survey

Our data were collected during the Ghana Cocoa Farmers Survey (GCFS). The GCFS was conducted between 2002 and 2010, jointly by researchers from the Centre for the Study of African Economies (CSAE), the Ghana Cocoa Board (COCOBOD) and the Ghana Statistical Service (GSS). The survey had more objectives. Primary one was creating evaluation of the Cocoa Abrabopa program. This program of micro financing for inputs was done in separate villages then we use in our panel, the farmers from Abrabopa program villages will not be used, because they cannot be considered part of representative sample of cocoa farming population. Second objective was to create representative panel of cocoa farming population of Ashanti, Brong Ahafo, and Western regions and to use this panel as control group for the first objective. For this objective different villages were selected and we will only use these in our research (Zeitlin (2008)). Combining these goals let to compromise in sampling in later years for the programme as will be described in the next section.

Some many questions were only added to the questionnaires in later years; some on the other hand were not asked anymore, this result in unavailability of some of the variables in some years. Moreover, some changes in questionnaires were made on the way during the data collection process in the respective year; however this will not be the case of the variables we will use for our research. (CSAE and COCOBOD (2006))

2.3.1. Sampling

In 2002, the sample consisted of total 492 households. This sample was designed as a random selection of cocoa-farming households in 25 villages. These villages were randomly selected with probability equal to their cocoa farming populations from the 1998/1999 Ghana Living Standards Survey (GLSS). Not all households drawn based on the GLSS could be traced, in such case another farmer in the same village was randomly selected from a listing of cocoa-farming households. Moreover, sampling unit in the second stage of selecting was not the household but the farmer, so multiple individuals in one household could be selected, but occurrence of this issue was rare (CSAE and COCOBOD (2006)).

In the later years new households were added either as replacement of household that could not be traced or as an addition to the sample. The replacement of households was done in two ways. First, if new owner(s) or cultivator(s) of the land could be found, the one using most of the land was selected as the replacement. Second, additional households were added to the sample, the way they were selected differs in different periods of the survey (CSAE and COCOBOD (2006)).

In 2004, due to time constraint and deviations from the sampling protocol, the additional households were draw in more different ways and cannot be considered random sample. However, for the purpose of our thesis we will consider the assumptions of random sampling to still be valid. This accounts for less than 10% of our dataset in the respective year (CSAE and COCOBOD (2006)).

In 2006, random selection of farmers from one new village was added from the Western Region. Additionally the replacing of missing farmers was done again, this time as a random selection of farmers that migrated from other villages (CSAE and COCOBOD (2006)).

In 2008, larger number (more then 200) of new farmers was added. Both in new villages and current villages. This made the problem of aging of our sample i bit less intriguing, however this problem still occurs throughout the panel and the sample age is slowly diverging from the structure of the underlying population (Zeitlin (2008)).

Last, in 2010, only replacement of missing farmers was done again.

Chapter 3
Methodology

3.1. Efficiency

In the most basic microeconomic production model, producer uses N inputs ($\mathbf{x}$) to produce some particular amount of one or more output(s) ($\mathbf{q}$). In functional form written as $\mathbf{q} = f(\mathbf{x})$. But when allowing for sub-optimal output to be produced we can consider production to be set by inequality $\mathbf{q} \leq f(\mathbf{x})$.

Given some $\mathbf{q}$ and $\mathbf{x}$, any difference between the produced output and f($\mathbf{x}$) can be considered technical inefficiency. Technical efficiency (TE) is basically the ratio between produced output (given $\mathbf{y}$ and $\mathbf{x}$) and the potential output, thus it can exhibit values in between 0 and 1, where 1 is firm that is being completely efficient, i.e. lies on the production frontier. Precise formulas defining TE differ with the orientation we use; this will be explained later in this chapter.

The technical efficiency, together with the technology and scale efficiency (SE) are the factors that compose the total productivity.

In theory, all firms should be efficient. However, different environment, skills or possibilities, can cause producers and firms to operate under the production possibility curve, so to be technically inefficient. Studing this issue is quite common among economists and many studies interested in efficiency on many different samples have been made.

3.1.1. Obtaining Efficiency Scores - DEA and SFA

There are two principal methods used to obtain the efficiency scores of the individual decision making units (DMU).

First method is the data envelopment analysis (DEA) is a non-parametric mathematical programming method. In this method basically the DMUs are compared with each other by solving minimisation (or maximisation) problem to find the most efficient ones and to obtain the efficiency scores. This method will be described more in detail later in this thesis.

The advantage of the DEA are that it does not require us to make any assumptions about the functional form of the production frontier (production function) and neither about the distribution of the efficiency scores. The down side of the DEA is that outliers and any errors in data directly influence our results.

The second method used is the stochastic frontier analysis (SFA). SFA is an econometric (parametric) method used to obtain production frontier and at the same time, like DEA, enables one to obtain efficiency scores of the individuals in the sample. One has to assume the distributional form which the production frontier follows. While SFA can have many different specification depending on the functional form assumed, in the simplest form, when linear production frontier is assumed, the model would be defined by the following equation:

$$q_i = \mathbf{x_i}\beta + v_i - u_i \tag{3.1}$$

where v_i is an error term with a normal distribution, and u_i is one-sided inefficiency effect, for which we have to make an assumption about its distribution, trounced-normal or half-normal are the most common used ones. Maximum likelihood estimation then can be used to solve this problem. For efficiency analysis in agriculture, SFA is nowadays used perhaps more often than DEA. It is mainly preferred because its big advantage over DEA, which is that it can account for statistical noise and obtain t- and F-statistics like during any econometric estimation. However, assumptions about the functional form of the production frontier and especially the distribution of the efficiency scores must be made. Misspecification of the functional forms can cause results to be invalid, especially when looking for the determinants of efficiency. Moreover, there is no way to test whether our assumptions are right or wrong. In this thesis we will favour using the DEA over the SFA, since similar studies, like Richman (2010) (on the earlier years of the same dataset) or Amos (2007) used SFA to perform efficiency analysis of cocoa farmers. Using the DEA and avoiding assuming the functional forms could also allow us to explore the problem deeper while taking a look at the first-differences of the efficiency.

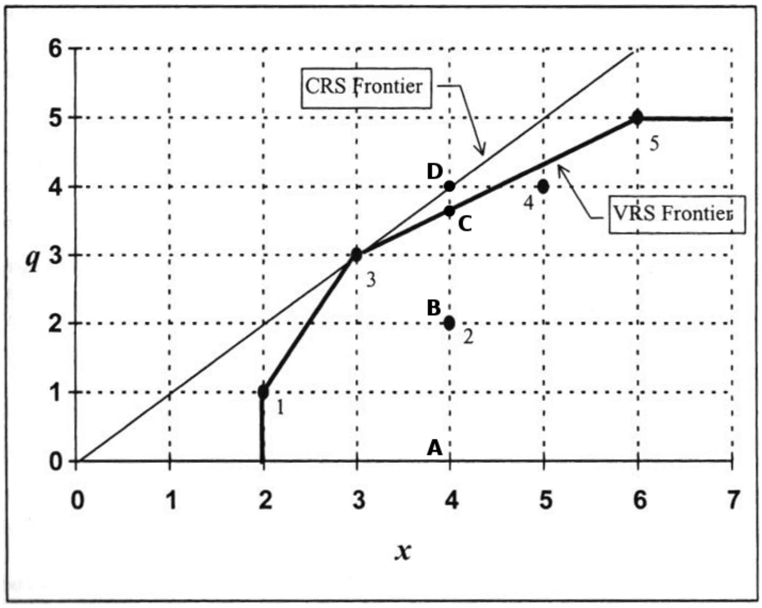
3.2. Data Envelopement Analysis

Data envelopment analysis in its earliest, input-oriented with constant returns to scale, form was initially introduced by Charnes et al. (1978), while many other specifications of DEA ware developed later.

3.2.1. Constant and Variable Returns to Scale

There are two basic specifications of the DEA problem. First is constant returns to scale (CRS) specification and it assumes the production frontier to exhibit constant returns to scales, the other - variable returns to scale (VRS) specification relaxes this assumption. Other specifications, like non-increasing returns to scale can also be formulated, however they are used less often then VRS and CRS ones. A imple example showing the CRS and the VRS frontiers is shown on the Figure 3.1. The points 1 to 5 illustrate a sample of observations. In this simplest single-input single-output example, CRS frontier is defined principally as a straight line passing through the origin and at last one observation in such way that all observations lay on the line or below it. The VRS frontier is combination of linear combinations of observations and straight lines parallel with axes in case no other linear combination can be made. And it is again constructed in a way so all the observations would lay on this frontier or below it.

Figure 3.1.: Example of CRS and VRS frontiers



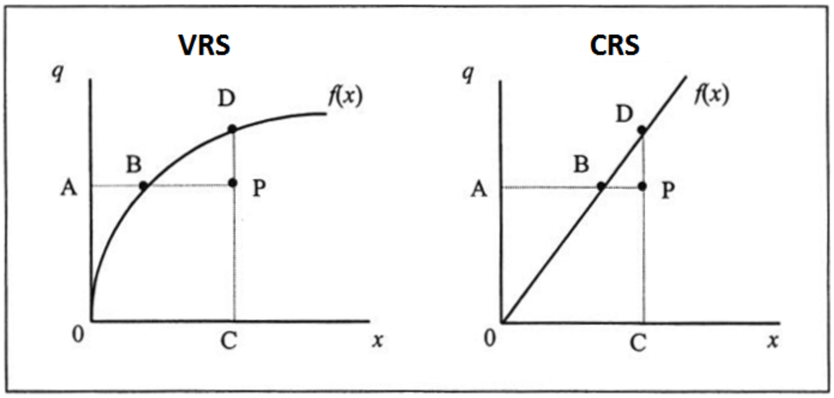
Source: Coelli et al. (2005)

As Coelli et al. (2005) explains, the CRS assumption is valid when producers operate at an optimal scale, but facing some unfavourable environment, like imperfect competition, government regulations, financial constraints or other, can cause them not to operate on the optimal scale. Especially due to the financial constraints and possibly also various other reasons, constant returns to scale are rather unlikely to be observed in our sample, therefore this thesis will use the variable returns to scale specification. This is also the most common choice among studies using DEA in agriculture, like Bravo-Ureta and Evenson (1994) or Coelli et al. (2002).

3.2.2. Output- and Input-Orientation

Efficiency can be measured in two ways, by using either input- or output-orientation.

Figure 3.2.: Input- and output-oriented technical efficiency measures



Source: Coelli et al. (2005)

In the input orientation, output is hold fixed and the inefficiency (1 - efficiency) is measured as by how much relatively can the amount of inputs be decreased to achieve the same output(s). When output-orientation is used, however, the inputs are fixed and inefficiency is measured as the proportional increase of output(s) that could be achieved. For the simple single-input single-output example showed on the Figure 3.2.

The input-orientation TE scores are counted as

$$TE = \frac{AB}{AP} \tag{3.2}$$

whereas the output-orientation are calculated as

$$TE = \frac{CP}{CD} \tag{3.3}$$

In the Figure 3.2 we can also see that when CRS are observed, the input- and output-oriented DEA models yield the same efficiency score. Moreover, both measures also identify the same set of DMUs as efficient, only for non-efficient DMUs the scores can differ when VRS are observed.

3.2.3. Model Definition

The linear programming maximisation problem used to obtain efficiency scores, with output-orientation and variable returns to scale can be defined as equation (3.4) (Coelli et al. (2005))

$$max_{\phi,\lambda}\phi, \tag{3.4}$$

$$\begin{aligned} subject\ to\ & -\phi\mathbf{q}_i + \mathbf{Q}\lambda \geq 0 \\ & \mathbf{x_i} - \mathbf{X}\lambda \geq 0 \\ & \mathbf{I1}'\lambda = 1 \\ & \lambda \geq 0 \end{aligned}$$

Where I denotes the number of the producers, N of the inputs and M of the outputs. For the i-th producer the inputs will be represented by column vector $\mathbf{x_i}$ and outputs by column vector $\mathbf{q_i}$. Data for all producers are represented by $\mathbf{X}$ the $N \times I$ input matrix and $\mathbf{Q}$ the $M \times I$ output matrix. Parameter λ is $I \times 1$ vector of constants. And where $\mathbf{I1}'\lambda = 1$ is a covexity constraint enabling DMUs to be compared only with others of its size and making the VRS specification, not allowing the linear combination with the origin, $\mathbf{I1}$ is na $I \times 1$ vector of ones. Parameter ϕ is scalar and $1 \leq \phi < \infty$, $(\phi - 1)$ is the propotional increase in outputs that could be achieved by i-th producer with current input quantities.

Technical efficiency score of i-th producer is then denoted as

$$TE = \frac{1}{\phi} \tag{3.5}$$

This problem must be solved I times to obtain score for each firm and it can be interpreted as that it takes the i-th producer and tries to radially expand its output vector $\mathbf{q_i}$ as much as possible while not exceeding the possible output set. In our single-output scheme, it simply maximises the expansion of the output given the coinstraint constructed as a linear combination of the values of other DMUs, that it can never exceed.

This single example of this linear programming maximisation problem would be denoted as folloving:

$$max_{\phi,\lambda}\phi, \tag{3.6}$$

$$\begin{aligned} subject\ to \quad & \phi q_i \leq \lambda_1 q_1 + \ldots + \lambda_I q_I \\ & x_i \geq \lambda_1 x_1 + \ldots + \lambda_I x_I \\ & \lambda_1 + \ldots + \lambda_I = 1 \end{aligned}$$

Where λ and ϕ are parameters, q_i is output of i-th observation and x_i its input. The first condition means that the expanded output ϕq_i cannot exceed the linear combination of other observations and the second condition restricts the maximization problem to situations with same or lower amount of inputs only. The third condition is, again, making sure only observations with comparable amount of input used are compared.

3.2.4. The Two Stage Approach

To examine the determinates of the efficiency, we will use the two stage approach, as it is recommended by Coelli et al. (2005) and moreover, it is used in the vast majority of the efficiency studies, including for example, Helfand and Levine (2004) and Haji (2007).

First Stage

In the first stage, the efficiency scores will be obtained by solving the DEA problem as described in equation (3.4). Usage of DEA on panel data is not so very common, especially in studies interested in agriculture. This thesis chooses to use observations from all years to serve as potential peers (i.e. to be compared with) for the entire sample, (expect the year 2008, which will only be allowed to serve as peers for itself, since there are some issues with capital, as will be explained later). Same approach was used for example by Biørn et al. (2003) in efficiency study of hospitals.

This way was decided, so the score would be comparable between the years. Because the sample partially differs in each year, otherwise the scores would be hard to compere between years and efficiency change in time impossible to examine and the advantage of panel data would not be possible to be used. SInce, as explained in Coelli et al. (2005), comparing efficiency scores from different samples is, at best, problematic.

The choice of inputs for the DEA problem this thesis uses is based upon the common classification as proposed for example by Coelli et al. (2005). The factors of productions used will be capital, labour, material inputs and land devoted to cocoa. This differs from the basic classification in two ways. First the land devoted to cocoa will be counted separately from other inputs, since it is very important itself in the production and also it cannot be in considered substitute to other inputs, moreover it would be difficult to valuate it to sum up with other inputs. Second, used services will be summed up with inputs, because the only service the farmers in our sample use (or we have data about its usage) is help from the government spray gang. Making one category of inputs and used services is a common approach supported by Coelli et al. (2005), with cocoa being the single output in our DEA model. Since the questionnaire was constructed with clear goal of researching the cocoa production, insufficient data is available about production of other corps by the farmers, so no other outputs will be considered, it will only be attempted to correct for them at least a little in the second stage.

Integrating all the different factors of production into this lower number is also desirable, otherwise including too many inputs in the DEA problem would result in too many DMUs appearing efficient and high mean efficiency scores (Coelli et al. (2005)). To solve the first stage DEA problem, the Open Source DEA (OSDEA (2014)) programme will be used.

Second Stage

In the second stage, the TE scores obtained in the first stage are to be regressed upon the environmental variables. Standard econometric methods can be used, so the direction and significance of effects can be tested for using the standard hypothesis. However, the disadvantage is that in the case when the variables used in first stage as inputs would be (highly) correlated with the environmental variables we use in the second-stage, the results would likely be biased. To address this issue, the variables from first stage will be also included in the second stage together with the new ones. To regress the variables and obtain efficiency scores, one can either use Tobit regression or standard OLS. McDonald (2009) argues OLS should be used, because efficiencies are neither censored data nor a corner solution and unlike other more advanced methods OLS can be performed by basic statistical software. This view is opposed by Simar and Wilson (2011), they explain that in this case OLS is consistent only under very specific assumptions and thus Tobit should be used rather than OLS. Moreover, in their other study (Simar and Wilson (2007)) they suggest bootstrapping methods to be used in the second stage. These methods are looking for the underlying data generating process and Simar and Wilson (2007) argue that this method yields valid inference in and thus should be preferred, they also support their claim with monte carlo evidence. Bootstrapping methods are, however, very advanced and they are well beyond the scope of this thesis. This thesis chooses to use Tobit regression, with censoring at 0 and 1. This is also supported as the method to use for normal applied economists by Coelli et al. (2005). Since this thesis can enjoy the advantage of panel data, the random effects version of Tobit regression will be used. The model can be, in the simplest way, notated as following: [1]

$$TE_{it} = \beta_0 + \mathbf{x_{it}}\beta + u_{it} \tag{3.7}$$

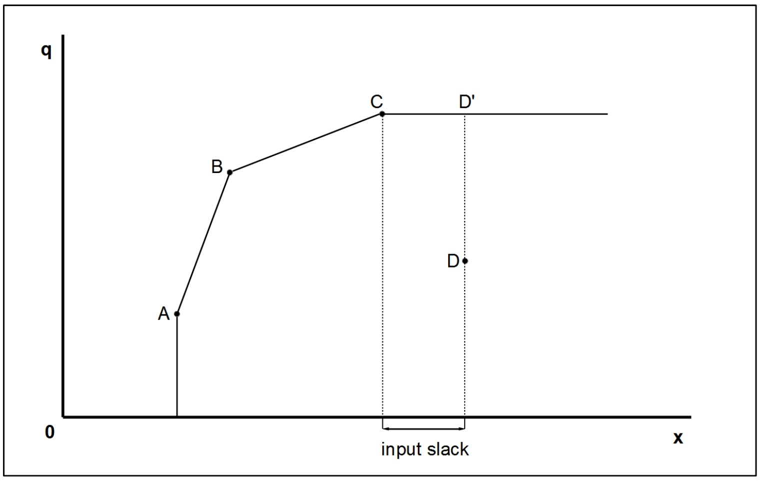
Tobit is also used by majority of other studies on efficiency using two stage method, both with DEA and SFA, Dhungana et al. (2004) and Richman (2010) to state some examples.

3.2.5. Input Slacks

Slacks might occur when the piece-wise linear frontier is running parallel with the axes. When the DMU is compared to such level of output, input slack occurs because the amount of inputs could be further lowered with still producing the same output. This would make the particular producers seem more efficient than they actually are. This issue is illustrated in the Figure 3.3 on the easiest single-input single-output output oriented example. There we can see that the producer D and its projection on the efficient frontier D', could be moved to the left, maintaining the same output **q** and while lowering the amount of input **x**.

[1] Where **x** is set of explanatory variables and it can also be logarithmic transformation of some variable, as will be case for the factors of production.

Figure 3.3.: Input slack example



Coelli et al. (2005) states that the issue of slacks is often overstated. Ferrier and Lovell (1990) argue that slacks are not really a technical inefficiency but rather the allocative inefficiency. We will accept this premise and since researching the allocative efficiency of our sample is not the goal of this thesis, we will not treat slacks in any way even though some more advanced methods have been proposed to solve this issue.

3.2.6. Scale Efficiency

Scale efficiency explains basically how much more the producer could produce, if he would operate on the optimal scale.

To precisely define the scale efficiency measure, we need to first define constant returns to scale DEA model. This one is basically the same as the VRS one, only relaxes the convexity constraint $\mathbf{I1}'\lambda = 1$. The linear programming problem for CRS output-oriented model is defined as in formula 3.8 (Coelli et al. (2005))

$$max_{\phi,\lambda}\phi, \tag{3.8}$$

$$\begin{aligned} subject\ to\ -\phi\mathbf{q}_i + \mathbf{Q}\lambda &\geq 0 \\ \mathbf{x_i} - \mathbf{X}\lambda &\geq 0 \\ \lambda &\geq 0 \end{aligned}$$

On the simple single-input single-output example in Figure3.1 the TE scores of the DMU ”2” for VRS are counted as

$$TE_{VRS} = \frac{AB}{AC} \tag{3.9}$$

and for CRS as

$$TE_{CRS} = \frac{AB}{AD} \tag{3.10}$$

.

The scale efficiency is defined as

$$SE = \frac{AC}{AD} \tag{3.11}$$

and because

$$\frac{AB}{AD} = \frac{AB}{AC} \times \frac{AC}{AD} \tag{3.12}$$

the following formula can be derived to define scale efficiency.

$$TE_{CRS} = TE_{VRS} \times SE \tag{3.13}$$

First we will obtain first both the CRS (by solving linear programming problem 3.4) and VRS (by solving linear programming problem 3.8) DEA models. Next we will divide TE_{CRS} with TE_{VRS} to obtain scale efficiency score. The scale efficiency will then be analysed using the two stage method; the same as in the case of TE. Same approach was used also for example by Dhungana et al. (2004) in the study of efficiency of Nepalese rice farmers. However, finding the determinants of the SE scores in more difficult and more correlations with various factors of production used in the DEA model are more likely to occur. For these reasons this task is not very easy. Thus in this thesis SE will be researched rather just in a way to see whether the determinants support in direction the one of TE or not.

3.2.7. Efficiency Change

The change of efficiency score will be counted between the two closest time periods where the efficiency scores are available for the particular farmer. Therefore it might not necessarily be two consecutive time periods. This was done because there were already many observations with missing values. About 15% of the differences do not come from consecutive time periods. For the possible different effects this could cause we will correct by including both the year dummy variables and the number of periods the difference consist of dummy variables.

The differences in efficiency will then be regressed by both differences and by their absolute value. The first is done to possibly further support the findings of efficiency determinants found in the second stage regression. The latter is done to find whether there are some determinats of efficiency change. The model can be notated as following: [2]

$$\Delta TE_{it} = \beta_0 + (\Delta \mathbf{x_{ait}})\beta + \mathbf{x_{bit}}\alpha + u_{it} \tag{3.14}$$

The efficiency change can exhibit values in between -1 and 1, however, unlike the efficiency scores, these are not censored, and therefore OLS regression will be used instead of Tobit.

3.2.8. Bias of the DEA

One problem of DEA is that a bias is almost always present in the efficiency scores. It is because whenever DEA is used on a finite sample, the efficiency scores obtained are upward biased, because not the whole population is observed and there simply must be some units that are even more efficient than the observed peers. Moreover, when used on a dataset with data noise present, the scores might be significantly downward biased and also individual ones upward biased (Coelli et al. (2005)). However, this thesis is based on the belief that this bias will not affect the results of determinants found.

[2] Where $\mathbf{x_a}$ and $\mathbf{x_b}$ are sets of explanatory variables, one variable can be included in both sets.

Chapter 4
Data

4.1. Key Variables

This section will describe more in detail all the key variables used in our DEA model.

4.1.1. Cocoa Output

The cocoa output is measured in kilograms of cocoa beans produced and is the sole output in our model. Since the OSDEA programme does not allow for zero output, these few observations were not used. For about 10% of the observation in 2006 sample, the output was in extremely high values (millions on kg produced), in these cases the information about 2006 output from the 2008 sample was used, same was done for observations with missing values of output in 2006 and 2008 samples, if the information was gathered in later years. In the cases where yield was higher than the experimentally achieved one of 2500 kg per hectare and year (Teal and Vigneri (2004)), the amount of cocoa was cut to the value corresponding to a yield of 2500 kg per hectare. This was the case of four observations.

4.1.2. Land

This variable ”land” is defined as the land used for cocoa production in acres. Since for any year but 2010, the aggregate area is not available, the sum of areas of plots used for cocoa production was counted instead. This however unfortunately results in higher number of missing values and lower quality of this variable in general. The values of land were compared in between years and in case of clearly wrong values manually set to the most likely value, or dropped if such value could not be identified. Moreover, observations with 0 area for cocoa were also dropped.

4.1.3. Labour

The variable ”labour” was constructed as sum of different tasks on the cocoa field performed by different groups. The tasks include: land preparation, planting, maintenance and weeding, applying inputs and harvesting. The groups are dividied in two separable categories. The first one is family labour, which comprises of man, women and children. The other one is hired labour, which consists of labourers hired on by tasks basis, the ones hired on by day basis, annually hired ones and nnoboa[1] labourers. Additionally in 2006 category of the young adults is present for family labour. And in 2004 the by task and by days labourers are counted as one group. This composition of labour allows us to research the effects of different types of labour. But on the other hand, the labour variable is composed of some 30 variables, where in each question about number of persons and days worked was asked and this brings up a lot of mistakes in data and makes variable labour quite problematic. Observations where harvesting was 0 were deleted, because it is impossible to produce any amount of cocoa without even harvesting it.

4.1.4. Capital

The variable ”capital” is constructed as sum of value of spraying machines and value of tools.

In each year the number of particular capital was took and multiplied by its price. Whenever the farmer did not pay for the capital, the mean price in the year was used in order to be able to valuate this capital in some way. Furthermore when price was not available the mean price was used again. In 2010 price data on some tools were not available; data from 2008 and 2006 corrected for inflation[2] were used.

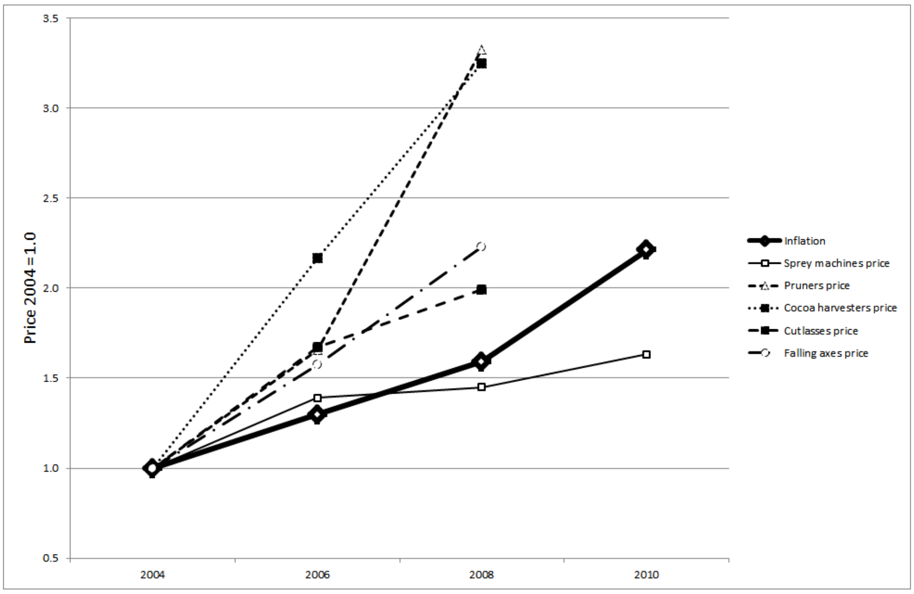
Tools composition differs in each year, at last in some period of time knapsacks, censor machines, pruners, cutlasses, drying mates, cocoa harvesters, falling axes, motorised spraying machines, hoes and other agricultural equipment were used. Moreover in 2008 only unsubstantial data about tools are available and tools are not counted at all, as mentioned earlier, year 2008 will not be used as peers for other periods. The

[1]Nnoboa is a local labour sharing scheme.
[2]World Bank (2013b) used as source of inflation throughout this thesis.

variations in capital will tried to be corrected for by using year dummy variables in the second stage analysis.

Finally the values of capital were set to 2010 price level on inflation basis. Since the composition differs slightly as already stated, using outside measure, like inflation, for correcting for price change rather than change in the mean values of the different types of capital was preferred. As we can see in the Figure 4.1[3], inflation is only somewhat similar as the changes of values of various tools and spray machines, it is quite lower then changes in prices of tools (the ones occurring in all 3 periods when tool prices are available), but higher than spray machines. Spray machines have however a larger share then individual tools in the capital. Therefore, inflation is considered somehow usable as proxy for the capital price change. The rest will hopefully be again corrected for by using year dummy variables in second stage regression.

Figure 4.1.: Mean capital prices and inflation



Last, for simplicity and unsubstantial data, the depreciation of capital is not considered.

4.1.5. Inputs

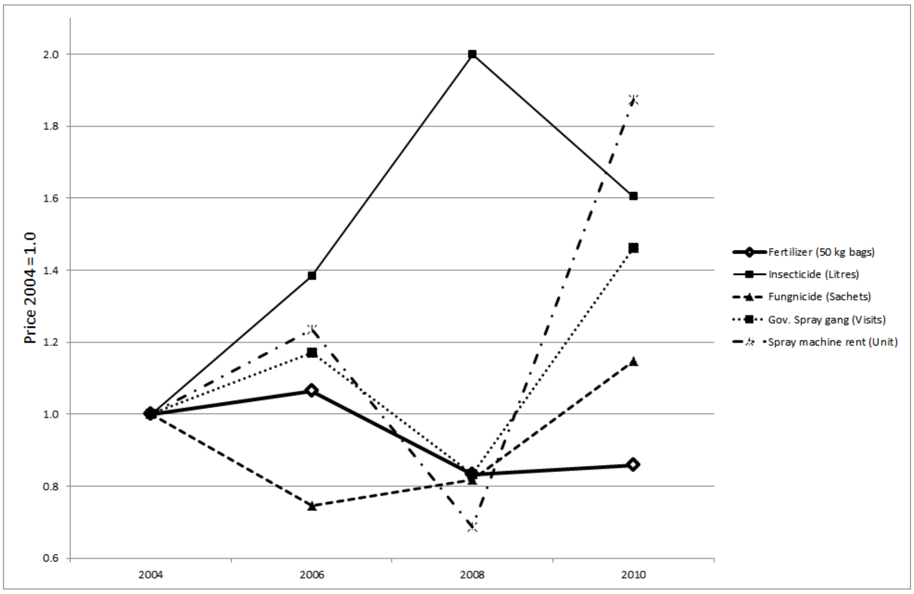
This variable consists of inputs used, services of government spray gang used and renting of spray machines. Inputs used are fertilizer, fungicide, insecticide and other inputs. For each the value was obtained as an multiplication of amount used by the

[3]For data about mean capital prices and inflation refare to Table A.1 in the appendix.

price per unit paid. Again where no prices were recorded or farmer heave not paid (quite common) the average price was taken (for the particular unit).

In case of inputs the price change in time is much less significant than in case of capital as we can see in the Figure 4.2[4]. Moreover price of fertilizer, which makes on average about half of the total value of the inputs, had got cheaper over the years. Therefore, the total value of inputs will not be in any way corrected for the price change over the years.

Figure 4.2.: Mean inputs prices and inflation



4.2. Data Issues

4.2.1. Missing Values

Our dataset generally exhibits high number of missing values. However, in cases of individual variables, the number of missing values is mostly quite small. Only when add up together this forms problem of a larger scale. If some variable exhibited larger number of missing values, it was used in analysis separately, as will be shown later in this thesis. In most cases missing values were not corrected for in any way. The observations were simply not used in the analysis.

The scale of the missing values problem is indeed quite large. No more than 2/3 of observations could be used for analysis of the efficiency determinants. In the

[4]For data about mean inputs prices refare to Table A.2 in the appendix.

case of analysing efficiency change this problem is even bigger, at best only 50% of observations were possible to be used. However, we believe the reason why some values are missing is uncorrelated with any of the variables used in our analysis, since as stated the problem is with high number of variables with some missing values, rather than few with a lot of missing. Furthermore, the number of observation we are left with in the end is sufficient to perform the desired analysis.

Observations with clearly wrongly entered data or further issues were treated same as the ones with missing values and not used in the analysis.

4.2.2. 2007 Ghanaian Cedi Redenomination

In December 2006 Bank of Ghana first announced that redenomination of Ghanaian Cedi will be made. The New Ghana Cedi was launched in July 2007 and co-circulated with the Old Cedi until January 2008 when it became the sole currency of Ghana. The conversion rate between the Old and the New Cedi is 1 : 10 000. (Dzokoto and Mensah (2010)). This redenomination causes sleight problems in our dataset. Some of the prices in the 2008 and 2010 sample are still denoted in the Old Cedis, these were converted to the New Cedis. Same all the prices until 2006, that are also obviously in the Old Cedis were converted to the New Cedis so the values would be comparable over the years.

4.3. Sample Description

Summary statistic for all the variables used can be found in the Table A.3 in the appendix.

It is important to note for all the individual characteristics of farmer, like age, education or gender, the characteristics of household head were used, even though the farms are usually family run.

4.3.1. Age

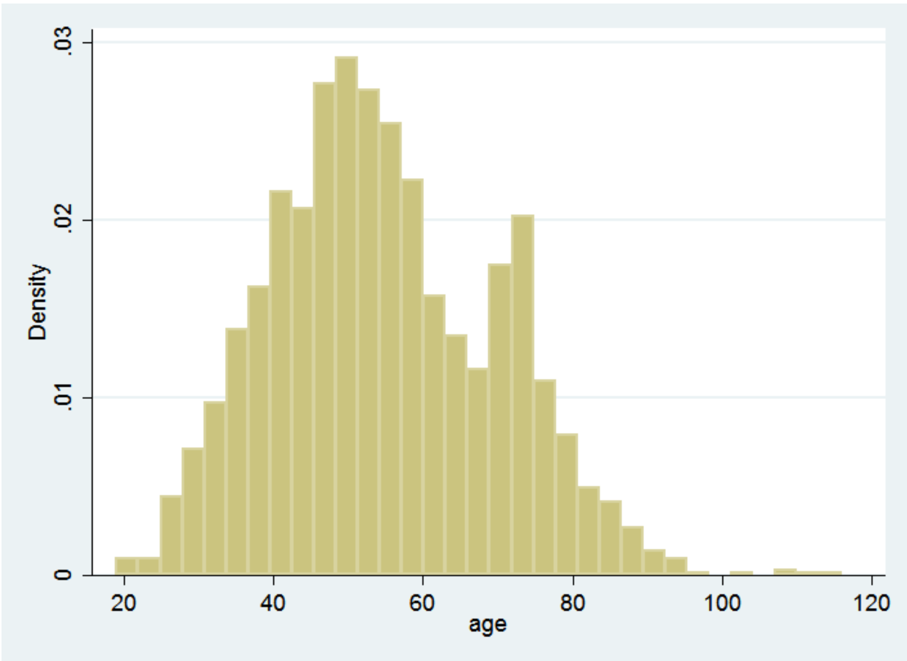
The mean age of the farmers in our sample is 54.8 years and the sample is slowly getting older, this is most obvious between the years 2008 and 2010, when no new farmers were added to the sample. However, there are no other great differences in age structure in the different years. The mean age in each time period of our sample can be seen in the Table 4.1.

Table 4.1.: Mean age in years

	2004	2006	2008	2010	all
Mean age	53.3	53.8	55.2	57.2	54.8

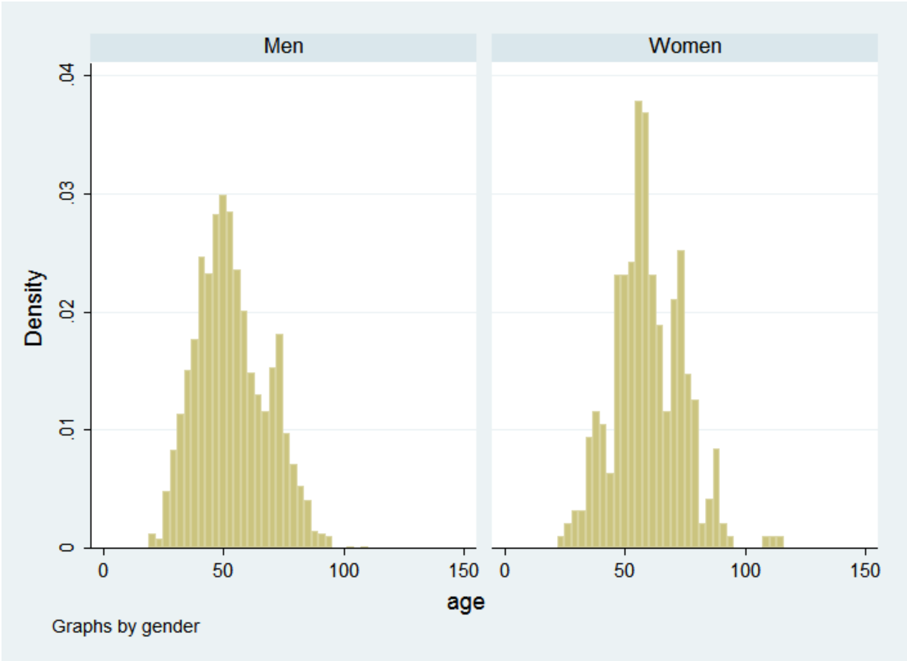
The age distribution as can be seen in Figure 4.3 in the sample is quite standard, perhaps, maybe some farmers are older then one would expect, but this is also caused by taking the characteristics of the family head.

Figure 4.3.: Farmers' age distribution



The age distribution among the genders is not the same, women are on average 59.72 while for man only 53.54. This could be caused both by the gender differences and by the fact that for older women it might be easier to act as a family head then for the younger ones.

Figure 4.4.: Age distribution by gender

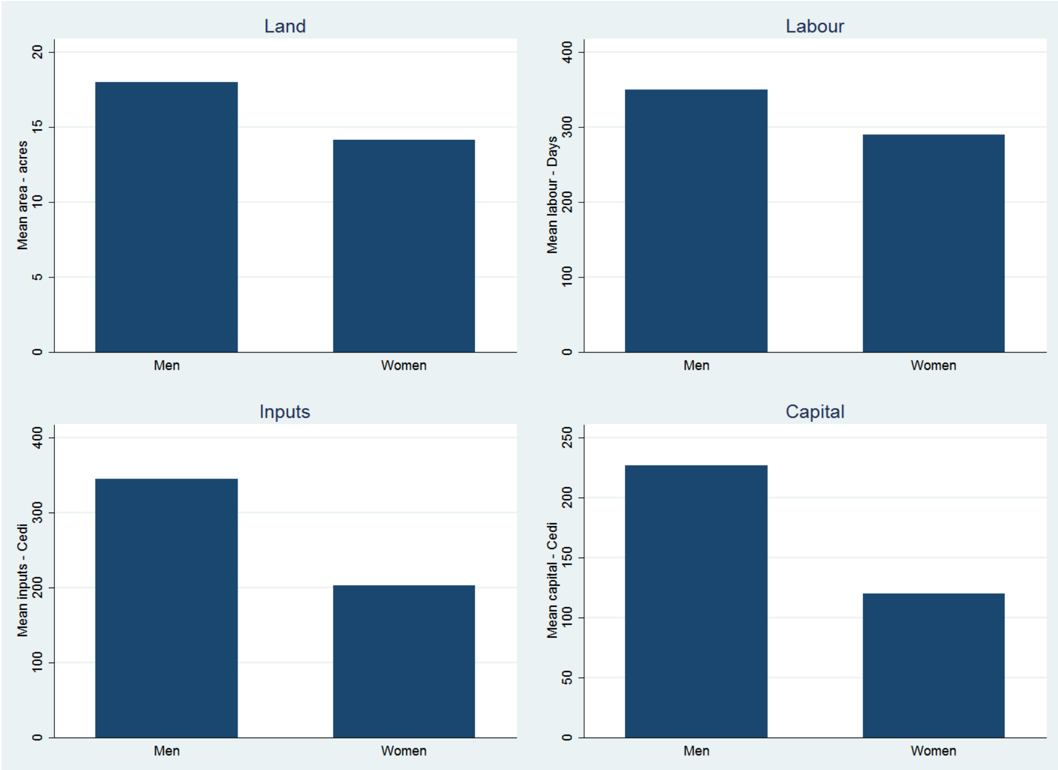


4.3.2. Gender

The gender distribution across our sample is far from equal. Only 16% of the farmers are women. Moreover, this share does not differ greatly over the years. This low share is likely caused by family arrangement, where women are less likely to be family heads. Moreover, this is supported by the fact that share of female household heads is much lower than their share in labour performed on farm, which is about one half of the labour done by men.

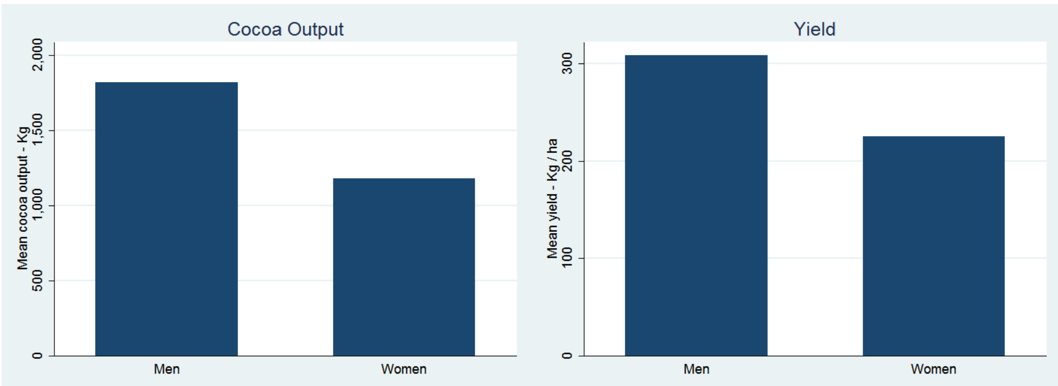
In the Figure 4.5 we can see that women use much less factors of production (simply inputs - as in our DEA problem) then men, in all categories, this situation is also consistent throughout the sample years. Moreover this difference is even higher for usage of non-labour inputs and capital that need to be purchased, unlike land and labour which is mainly performed by family. This may suggest female farmers could either face unfavourable conditions, or possibly they could have other, more important priorities then cocoa farming. However, the share of land devoted to cocoa does not differ between male and female farmers.

Figure 4.5.: Distribution of the factors of production by gender



Given the difference in amount of inputs, it comes as no surprise that the total amount of cocoa produced and yield per hectare are also lower for women then for men, as we can see in the Figure 4.6

Figure 4.6.: Cocoa output and yield by gender

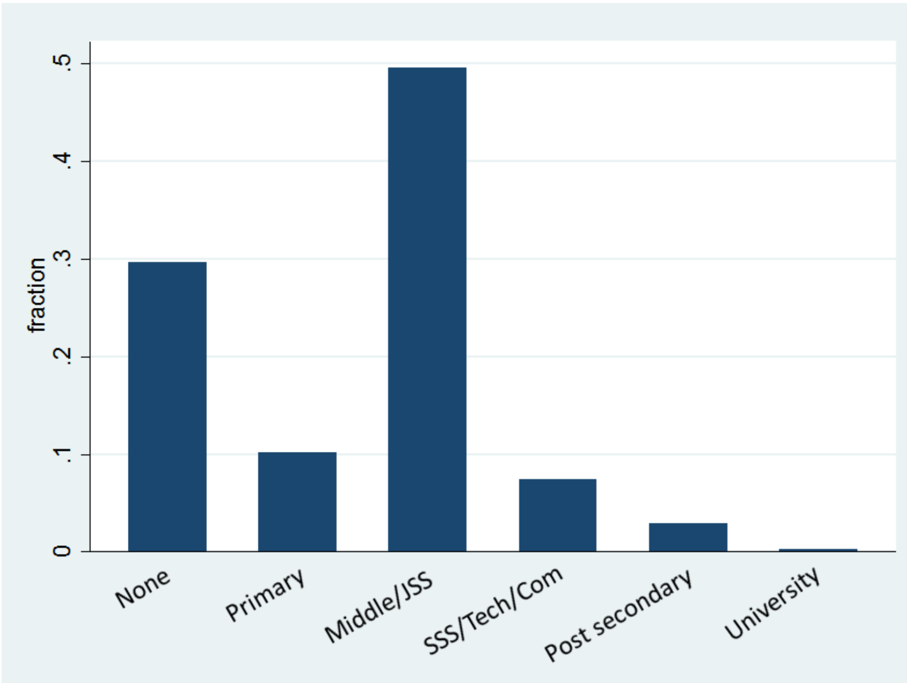


Because of the fact that factors of production and gender are correlated, we will have to include them in the second stage regression, otherwise the issue of relation of gender and efficiency could not be properly analysed.

4.3.3. Education

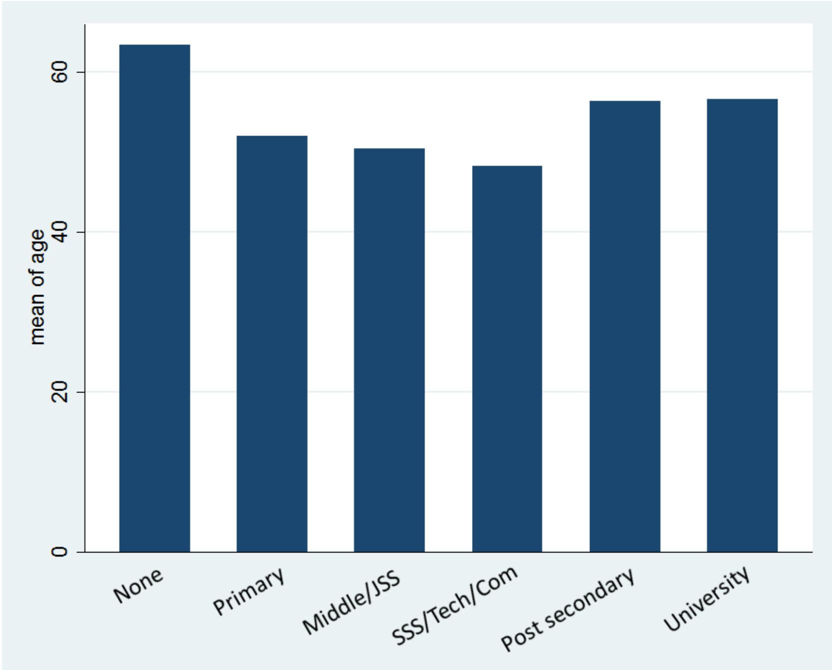
There are six levels of education used in our sample are: none (0), primary (1), middle / junior secondary school (2), senior secondary / technology / commerce school (3), post-secondary (4) and university (5). The highest level attended (not necessarily finished) was taken for each farmer. The mean education level in our sample was roughly 1.4 and stays almost stable throughout the yeas. Half of the farmers in our sample had middle / junior secondary school, the distribution of education can be seen on Figure 4.7 and 89.5% of all farmers had middle or lower education.

Figure 4.7.: Education distribution



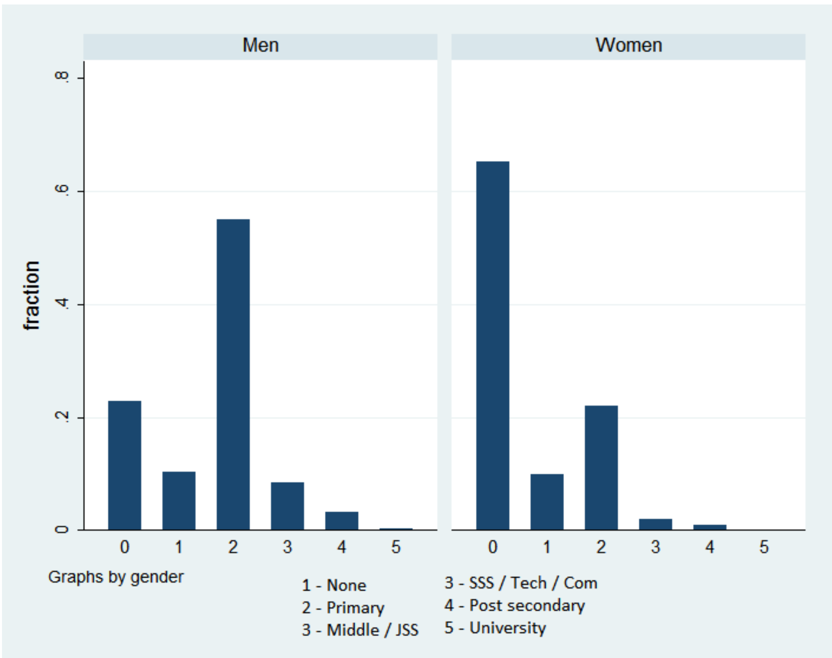
The mean age of the decreases with higher education in our sample as we can see in Figure 4.8, only for the two highest categories this is not the case, these two categories however have only smaller number of observations.

Figure 4.8.: Mean age for education levels



The Figure 4.9 shows that education is not equal for male and female farmers, male farmers are much more educated. The mean level of education for men is 1.6 while for women it is just about 0.6. This can partially be caused by the correlation with age and partially by social conditions at the time they could gain education.

Figure 4.9.: Education distribution by gender



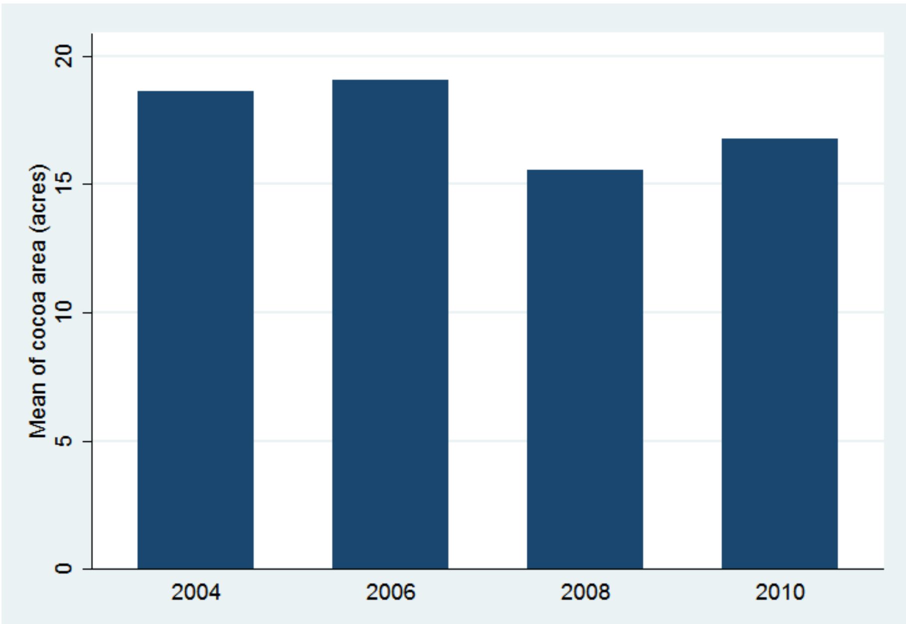
There are no clear relations between education and levels of factors of production used, nor with yield and cocoa production. Data rather suggest that the higher level of education the higher the usage, output and yield, however this relation does not seem to be necessarily the same over the time periods, especially for the higher levels of education with fewer observations.

4.3.4. Factors of Production and Output

Land

The Figure 4.10 suggests that mean area devoted to cocoa is rather decreasing over the years. Similar pattern also holds for mean of total land, and this data were gathered with different question, not summing up more values, thus it seems this observed pattern should not be caused by errors in the data.

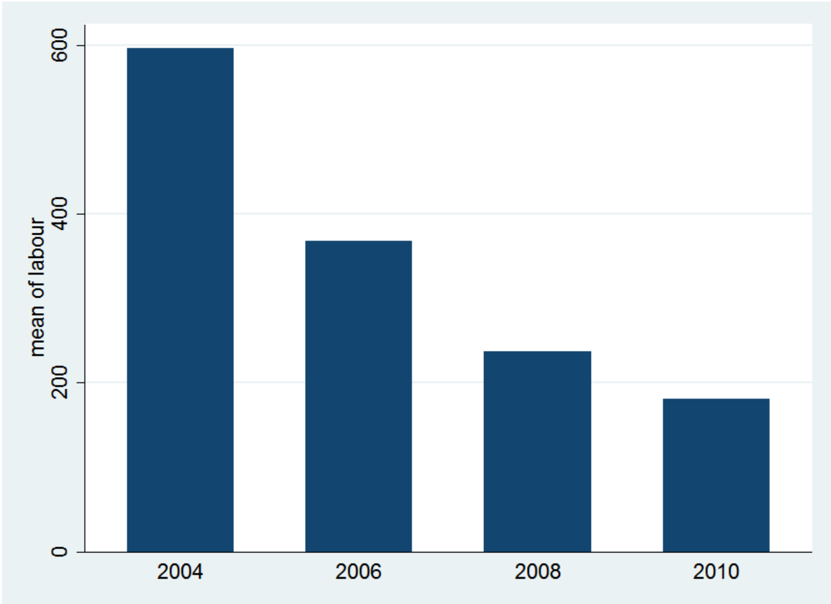
Figure 4.10.: Mean land for cocoa in years



Labour

The mean labour days performed on the farms to produce cocoa are decreasing rapidly throughout the sample, with the value in 2010 being less than one third of the value in 2004 as can be seen in the Figure 4.11.

Figure 4.11.: Mean labour days in years

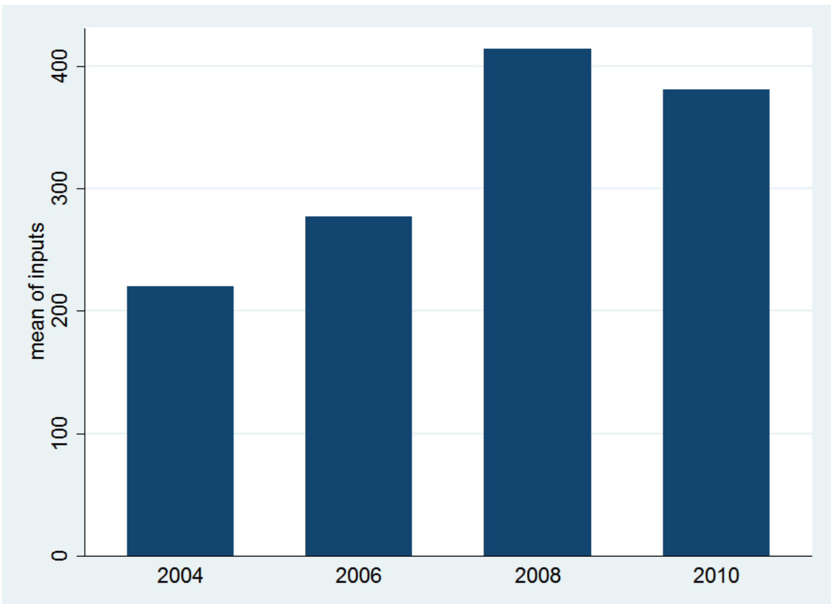


The mean number of days of household labour and hired labour are similar in our sample.

Inputs

As shown in the Figure 4.12 data suggest that the usage of inputs is increasing over the time. However, this can be influenced by price change, but which was not identified and corrected for.

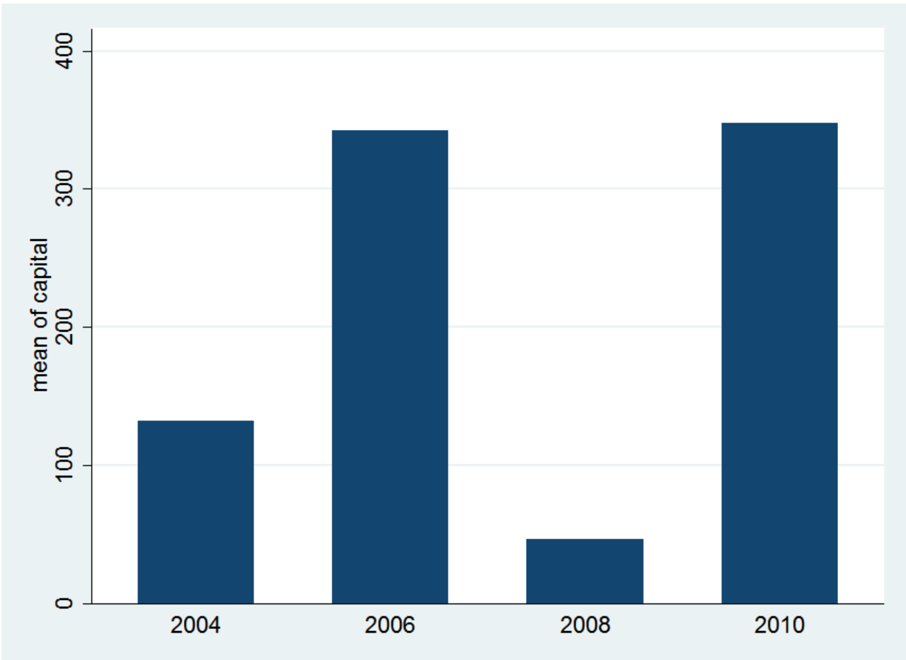
Figure 4.12.: Mean inputs value in years



Capital

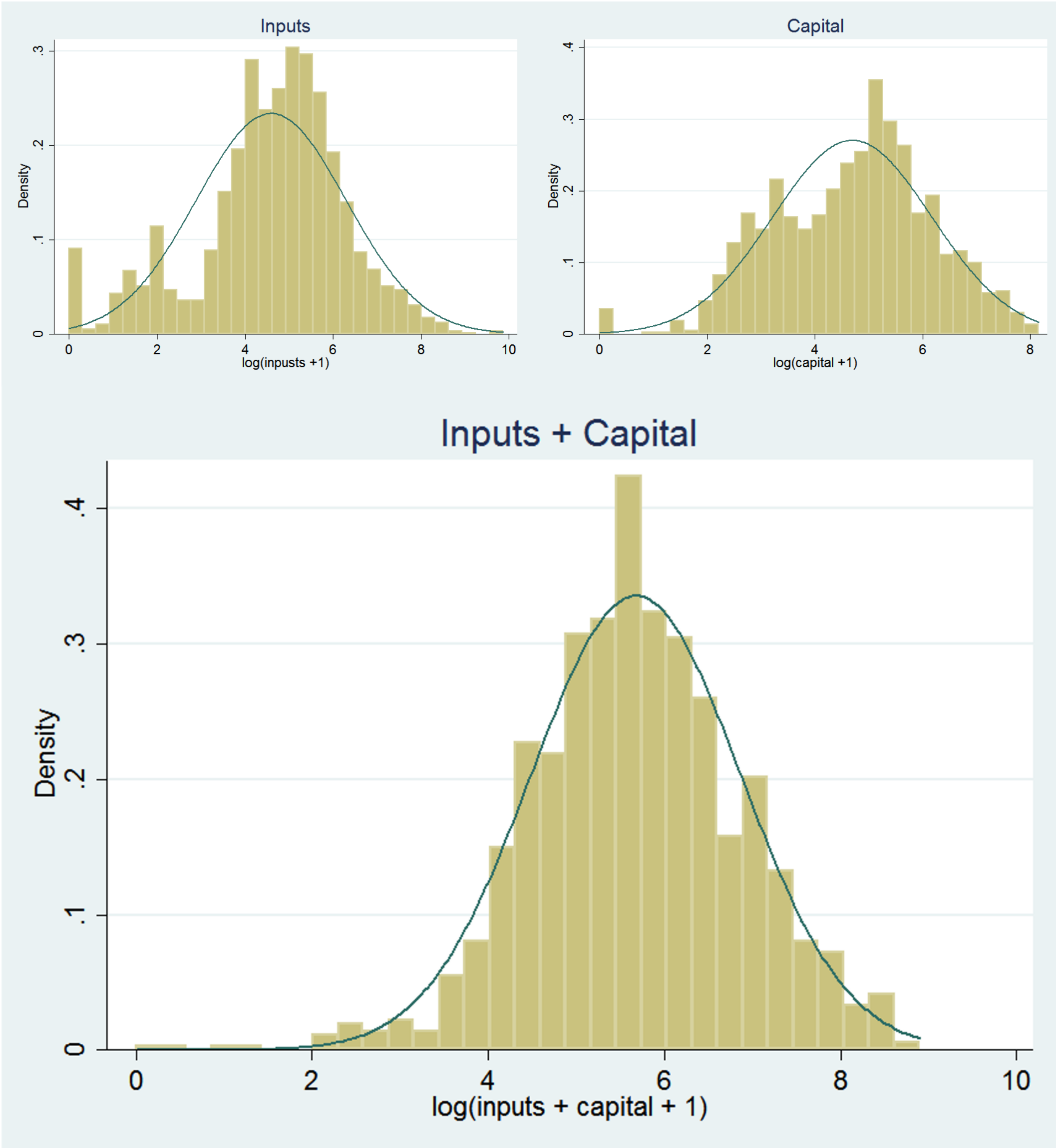
It is just a bit too difficult to judge anything about the change of capital in time, since as stated earlier the composition of tools courted to capital slightly changes over the years and more, the capital in 2008 is composed only of spraying machines, no tools. But from Figure 4.13 it rather seems that usage of capital is also increasing with time.

Figure 4.13.: Mean capital value in years



Moreover, for farmers in our sample inputs and capital are somewhat substitutes. This is because if spray machines are used they are either included in inputs as government spray gang visit or renting a pray machine or they are included in capital if farmer owns some. This is supported by Figure 4.14 where we can see this graphically.

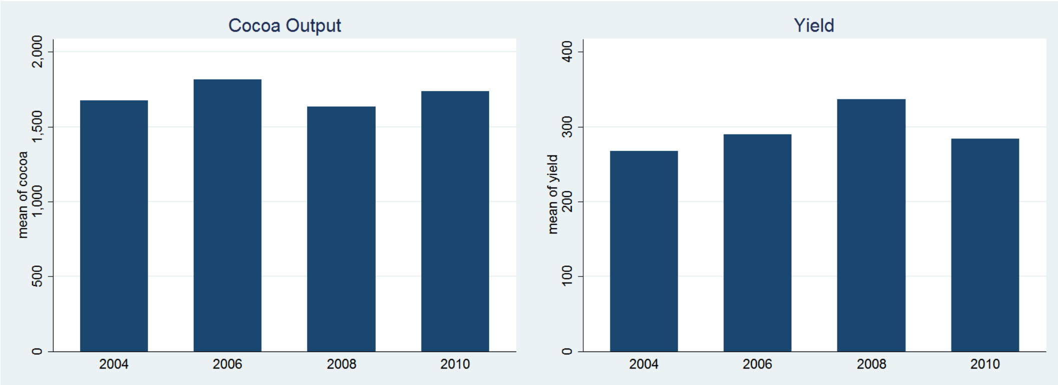
Figure 4.14.: Distribution of logarithms of inputs, capital and their sum



Output

The mean output of cocoa more or less stagnated over the time period of our sample as one can see in the Figure 4.15. The yield grew in the first three time periods but then decreased in 2010.

Figure 4.15.: Output and yield in years



Chapter 5

Results and Discussion

5.0.5. Efficiency Scores

The mean technical efficiency we found for the farmers in our sample is 35.4%. As can be seen in the Table 5.1 the mean TE score is slightly decreasing in time. However, the peers[1] are more or less equally distributed throughout all the time periods in our sample. The minimum of efficiency was about 1.1%. There are in total 92 observations found to be the efficient ones. This means the frontier is formed by about 6.3% of observations. Total of nine farmers serve as peers more than once - in more time periods, all those nine do exactly two times.

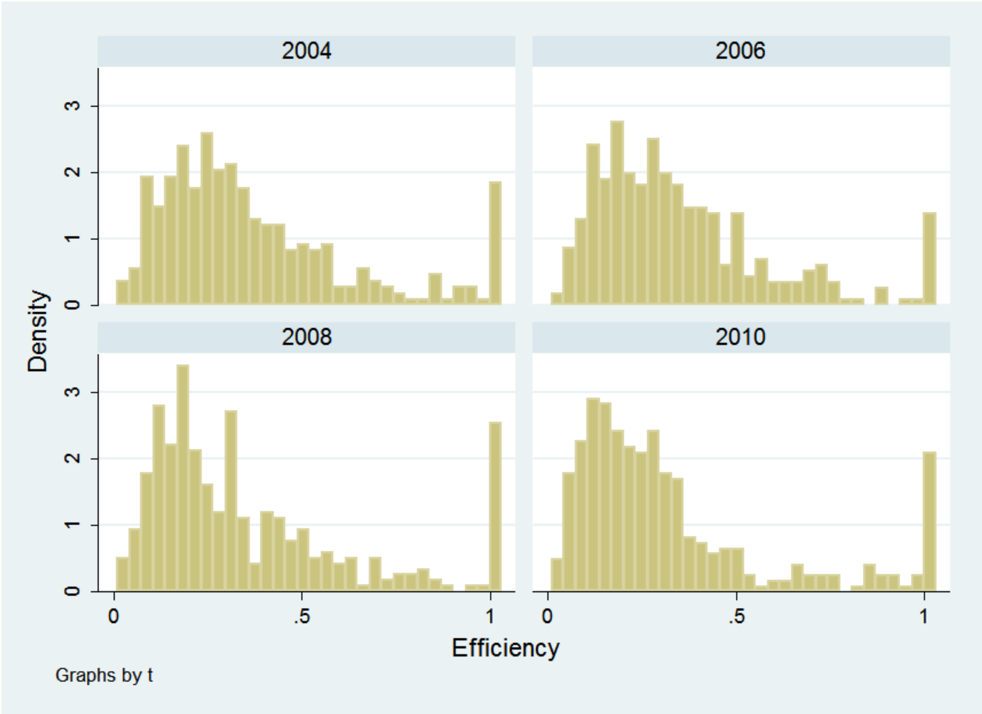
Table 5.1.: Summary statistics of the technical efficiency scores

Technical efficiency	n	mean	sd	min	max	nu. of peers
2004	339	0.376	0.254	0.011	1	20
2006	364	0.355	0.231	0.020	1	16
2008	370	0.360	0.269	0.024	1	30
2010	389	0.328	0.267	0.018	1	26
total	1462	0.354	0.256	0.011	1	92

The efficiency distribution in our sample also stays similar for all the years, as shown in the Figure 5.1.

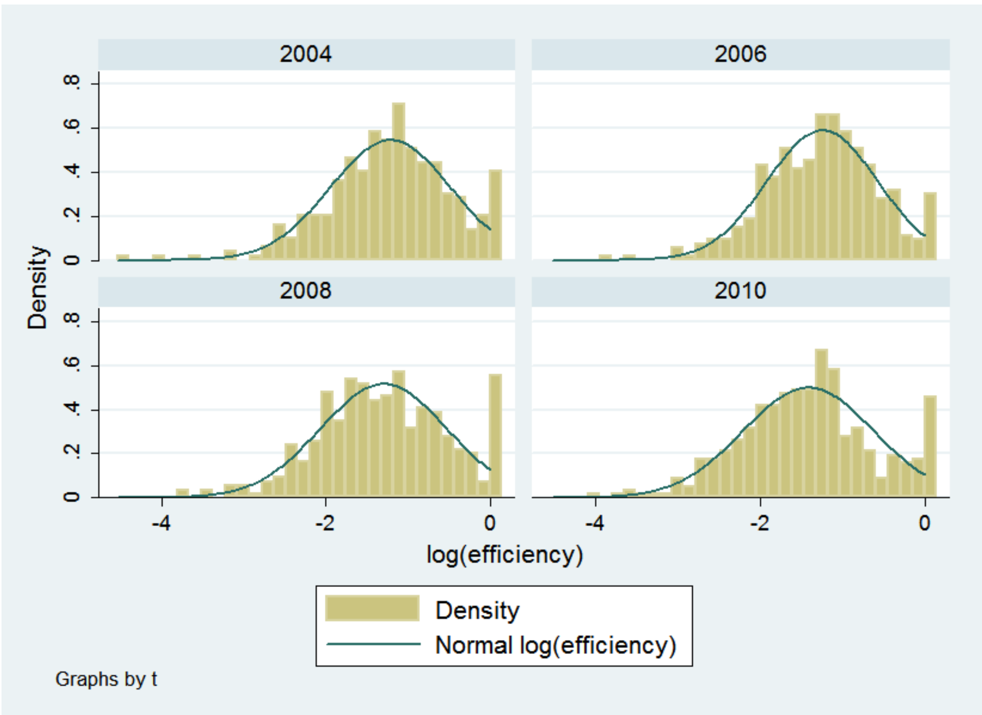
[1] Peer, i.e. the observation with TE = 1.

Figure 5.1.: Efficiency distribution in years



Furthermore, in the Figure 5.2 we can see that the efficiency distribution follows the log-normal distribution censored in 0.

Figure 5.2.: Logarithm of efficiency distribution in years



The Table 5.2 shows the shares of farmers in categories by efficiency, it should be noted that more then 60% of our farmers had efficiency between 10% and 40%. Moreover median of efficiency was just about 28%.

Table 5.2.: Efficiency distribution

Efficiency	n	share	cumulative
<0.1	131	0.090	0.090
0.1-0.2	355	0.243	0.332
0.2-0.3	296	0.202	0.535
0.3-0.4	229	0.157	0.692
0.4-0.5	135	0.092	0.784
0.5-0.6	90	0.062	0.845
0.6-0.7	47	0.032	0.878
0.7-0.8	40	0.027	0.905
0.8-0.9	28	0.019	0.924
0.9-1	111	0.076	1

Mean efficiency scores found by this thesis are substantially lower then the scores found in other studies concerning cocoa farmers. Richman (2010) and Danso-Abbeam et al. (2012) founded mean TE scores for Ghanaian cocoa farmers to be 44% and 49% respectively and Onumah et al. (2013) founded TE scores 85% for cocoa farmers in Eastern Ghana. Amos (2007) and Eyitayo et al. (2011) founded mean TE scores for Nigerian cocoa farmers to be 72% and 87% respectively. And Binam et al. (2008) in his meta-frontier analysis of cocoa producers in West Africa founded intra-country mean TE scores to be 44% in Ghana, 58% in Côte dIvoire, 65% in Cameron and 74% in Nigeria.

Eyitayo et al. (2011) same as this thesis used DEA, in his case to research TE of Nigerian farmers. However it is impossible to compare our TE scores with his for two reasons. First, his specification differs from the one used in this thesis, since he only uses land, cost of labour, fungicides and pesticides as the inputs. Second, his sample is much smaller, just 60 farmers, and thus he finds 68% of the farmers to be efficient. In the other study from Nigeria Amos (2007) used SFA, with half normal inefficiencies distribution. However, assuming thes distribution of inefficiencies is rather too restrictive and in most cases half normal distribution will yield higher mean TE score then trounced-normal distribution specification. And thus it is again difficult to compare the TE scores.

When trying to compare our results, it is probably best to do it by comparing with Richman (2010), Danso-Abbeam et al. (2012) and Onumah et al. (2013), who all used SFA with trounced-normal inefficiency distribution. It is important to note that Richman (2010) used the earlier years of the same dataset, so here we can consider this to be comparison of scores obtained by different methods on basically the same sample. Our scores are about 10% lower then the ones found in his study. Moreover, also Danso-Abbeam et al. (2012) finds higher TE scores then our study, in this case by 15%. These differences are likely to be result of a negative bias in our TE estimates caused by data noise or outliers that DEA could not cope with. The previous two results are also supported by findings of Binam et al. (2008), who too used SFA. The scores found by Onumah et al. (2013) are rather surprisingly height.

To conclude, it is very difficult to compare our mean TE score with other studies, but it seems reasonable to accept that we have a negative bias in the TE scores we have found. But still, we believe that this bias will not influence the findings inference during the analysis of determinants of efficiency and its change.

5.0.6. Efficiency Change

The mean efficiency change in our sample was close to zero, just about -1%. And the efficiency change scores ranged from -98% to +98%. The mean magnitude of change[2] was a quite high 21%. The variance in efficiency change scores increased significantly in 2010, refer to Table 5.3. It is partially caused by higher number of observations that are counted as difference over more periods than one, but even when counted without these, the variance in 2010 was still higher than in the two previous periods. The variance in efficiency change in our sample is in general higher then we would expect, and it could suggest the farmers' efficiency is not very stable for particular individuals over time, which could be quite important phenomenon.

Table 5.3.: Summary statistics of the efficiency change

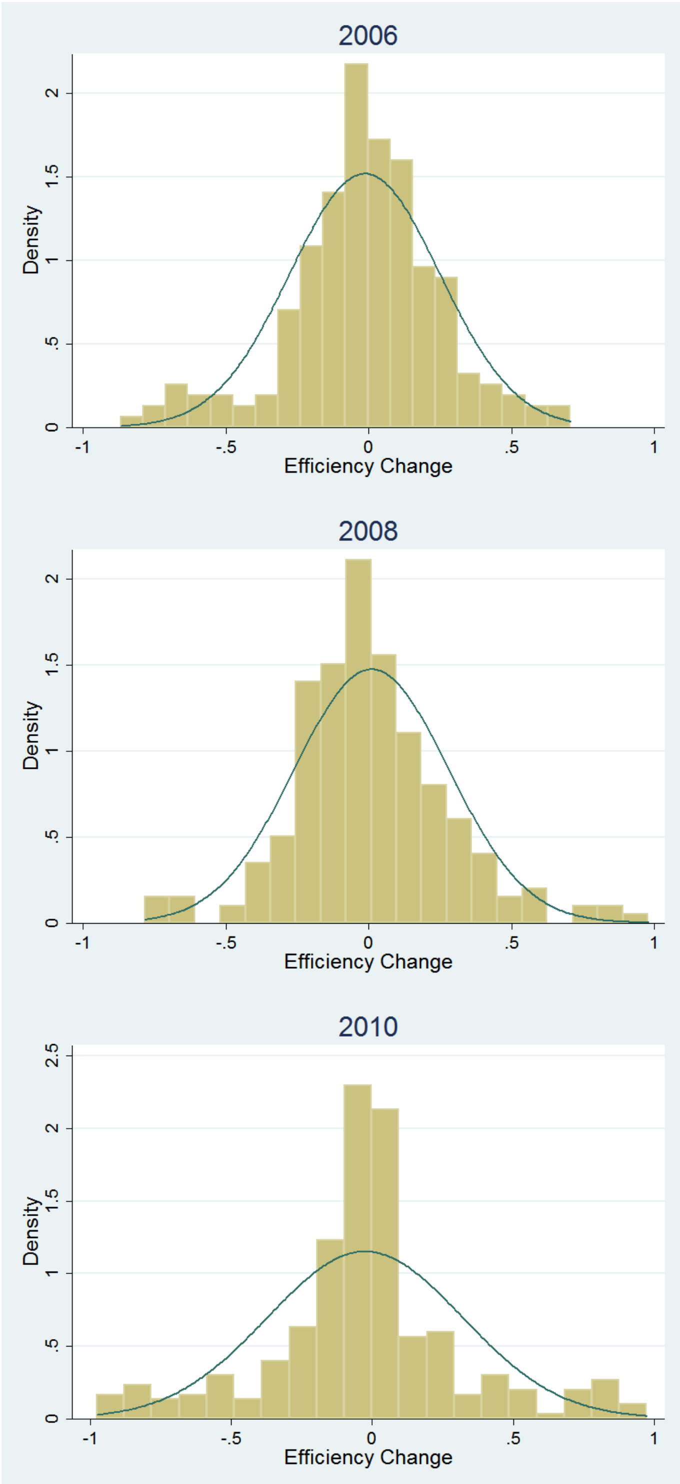
TE difference	n	mean	sd	min	max
2006	199	-0.0138	0.2629	-0.868	0.707
2008	226	0.0098	0.2707	-0.783	0.980
2010	308	-0.0256	0.3461	-0.976	0.975
total	733	-0.0115	0.3028	-0.976	0.980

In the Figure 5.3 we can see the histogram of the distribution of the efficiency change scores in the three time periods. We can see that the distribution in 2010

[2]Counted as mean of absolute values of efficiency change.

indeed quite differ from the ones in the previous two periods. This phenomenon is however difficult to explain, there is no obvious cause to this.

Figure 5.3.: Efficiency change distribution in years



5.1. Determinants of Efficiency

Total of fourteen of different random effects Tobit regressions were estimated to research the determinants of the technical efficiency. The results of regressions (1) to (5) can be seen in the Table 5.4 and the results of regressions (6) to (10) can be seen in the Table 5.5. The first three regressions are the basic model differing only in usage of different dummy variables. First regression (1) explaing the basic model includes variables for factors of production, farmers' characteristics, technical specifications, farm management style, investments and time dummy variables. The second regression (2) does not include time dummy variables. The third regression (3) includes extra village and regional specific dummies. The regression (10) is specification without usage of inputs dummies.

The regression (4) to (9) are performed to find effects of some other variables, that due to missing values or incomplete data issues could not be included in the basic model, especially together, since otherwise the number of observations would be greatly reduced. Fot the basic three regressions we have above 1300 observations, for the other less, even as few as 384 in regression (5), because it includes only data from 2010. The last four regressions performed in this way, including dummy for paing share of output to annual labourers, dummy for attendance in church or mosque, dummy for being a prominent and dummies for separate negative events, are not reported and they found no significant effect of these variables. All the determinants of efficiency are individually presented and discussed next. It is just now important to remind, and bear later in mind that mean TE found in our sample was just about 35% and median some 28%, so for example effect of determinant that is dummy variable of 10% on efficiency would be already quite huge.

5.1.1. Factors of Production

First, it is important to mention that the results of the influence factors of production have on TE can be less reliable, since these variables are already included in the DEA model. However, it is still important to discuss the results. Moreover, these variables were also included to prevent possible bias in estimates for variables correlated with them like age or gender.

Land Devoted to Cocoa

The land used to cocoa production (or its logarithm) is found to have very significant and positive effect on efficiency. And being significant at 99% confidence level in all three basic specifications of our model. This suggests that larger farms are more efficient. This finding is consistent with findings of Richman (2010) and Amos (2007) on samples of Ghanaian and Nigerian cocoa farmers respectively. Both studies also found the effect of farm size to be positive and significant. However this finding is not further supported by the findings of the effect of area change on the efficiency change, as can be seen in Table 5.6 no significant effect was found.

Labour

The estimated effect of labour usage on technical efficiency is found to be very significant and negative, with significance level over 99% in all of our specifications. Moreover, this findings are supported by the significant and negative findings of influence of difference of labour usage on efficiency change. This would suggest that labour is being overused by farmers in our sample. This can be supported by the high variance in levels of labour used.

However, any variable included both in second stage and in the DEA model is expected to have negative sign in the second stage, because the DEA model from its definition is trying to find other most efficient situation to compere our observation with and when increasing some input this efficient situation can only be more further from our particular observation and thus it makes it less efficient. But well, in the end, this can be considered to be over-usage of labour.

Also, the high variance is likely to be at last partially caused be the data noise resulting from the way this variable was constructed as stated earlier in the part about data of this thesis. If this issue would be of a larger scale, it would result in significant negative bias in the estimate of labour effect on the TE.

Capital

Capital was only found to have significant negative effect in the specification (2) when time dummies are not used an in the specification (9) researching effect of seed source, in all other the effect was not significant. The significance in specification without time dummy variables shoud not be given much importance, since as explained earlier the capital definition differs throughout the years and time dummy variables are important to be used in this case. Moreover, the same issues as for labour apply, since capital is also included in the DEA model itself. Therefore we can not conclude any clear effect of capital on efficiency or clear over-usage of capital.

Inputs

A significant negative effect of inputs usage on the TE was found in specifications (2) without the time dummy variables, (3) with village specific dummy variables and

(10) without the dummy variables for usage of inputs. Especially the last case tells us that, as it is quite obvious, there is correlation between inputs used and whether the farmers use individually some of the inputs. With same problems taken into account as with capital and labour, it seems there is rather over-usage of inputs in our sample. But still, again, it could be just caused by data noise.

5.1.2. Farmer Characteristics

Age

Age was found to have a significant negative effect, but only in the specification when no time dummy variables are used. However, this is likely caused by the fact that age is increasing in time and efficiency is decreasing over time. And even then this effect is with negligible real life consequence, difference in 10 years in age would make on average farmer only less than 1% less efficient. To conclude, no significant effect of age on TE is found. This is in accordance with other studies on cocoa farmers, where Danso-Abbeam et al. (2012) and Onumah et al. (2013) found no significant effect, both for Ghanaian cocoa farmers, Amos (2007) found negative effect for Nigerian farmers, and Richman (2010) finds positive effect for the farmers from same survey.

Gender

Our results suggest that female farmers are on average about 4% less efficient then males. However, it is important to note that this effect is significant just on 90% confidence level and in some specifications is not even found significant. Findings by Richman (2010) and Onumah et al. (2013) also suggest female cocoa farmers in Ghana are less efficient. Danso-Abbeam et al. (2012) found no effect of gender on efficiency. This suggests female farmers could be somewhat disadvantaged.

Education

Danso-Abbeam et al. (2012) and Onumah et al. (2013) found no significant effect of education on farmers' TE. However this thesis finds significant and positive effect of education on efficieny, same as Richman (2010) and Amos (2007). Depending on specification the average effect ranges between 1.5% and 2.5% for each extra educational level. Moreover, this effect is significant in all of our specifications in some even on 99% confidence level. Positive effect of education on efficiency have been found in many studies not just on cocoa farmers, some examples of other studies are the ones of Dhungana et al. (2004) on Nepalese rice farmers or Haji (2007) on mixed farms in Ethiopia.

Usage of Credit

One of the more interesting findings of this thesis is negative relationship between usage of credit for cocoa production and efficiency. Farmers that borrowed for cocoa production are on average about 3% less efficient, this effect is also significant in most of our specifications on either 90% or even 95% confidence level. Especially interesting is that these findings are even further supported by the estimated negative effect of change in credit usage on efficiency change as can be seen in the Table 5.6.

It is difficult to argue why there is this negative effect. Possibly it could be due to investments that only would yield results later and in short-run decrease efficiency. Or perhaps farmers using credit could be the less favoured ones facing tougher credit constraint that does not allow them to operate efficiently, as can be seen in Table A.4 in the Appendix, farmers using credit do dispose of slightly lower levels of capital and inputs. This explanation, however, is not so strong in explaining the effect on change in credit usage on efficiency change. Other possible explanation could be some unobserved effect differing for farmers that borrow and those who do not. Last possible explanation this thesis offers is much more speculative, it could be that credit usage is connected to stress and other effect that negatively influence farmers' performance. In any case, it would be interesting to design some research focusing more on this particular issue.

To compere other results, existing literature rather suggest positive effect of credit, but some studies also found negative effect. For example: Bravo-Ureta and Evenson (1994) on sample of Paraguayan peasants or Helfand and Levine (2004) on Brazilian farmers, found positive effects of credit on TE. But Haji (2007) on mixed farms in Ethiopia found negative effect of credit usage. Of the studies concerning cocoa production only Onumah et al. (2013) looked at the issue of credit and found positive effect of access to credit.

Savings

Farmers saving money for cocoa production were, unlike the ones borrowing, found to be on average about 4.5% more efficient. Significant in all specifications in some with over 99% confidence level.

Let us look at difference in levels of factors of production of framers who save for cocoa and those who do not as shown in Table A.4 in the Appendix, this shows that farmers who save have generally higher level of all factors of production, so it could be the farmers that are facing more favourable conditions allowing them to be more efficient. Or maybe the savings the farmers made just allow them to operate more efficiently. Perhaps it could also be the case that more capable farmers, that are more efficient, got themselves in better situation that allows them to save. No other literature researching this issue was found to be compared with.

Ownership of Land

Farmers owning their land were found more efficient. With the difference in TE between not owning any and owning all of the land that would be about 5.5%, this figure again differs with different specification and is found significant in most, with notable exception being usage of village dummy variables (3). This result is furthermore supported by the fact that difference in proportion of land own has positive effect on efficiency change and suggests that possibly, farmers' motivation could play positive role in determining their efficiency, as one would expect. This will be also briefly looked upon while researching some management variables and farmers' focus on cocoa production later in this chapter.

Other

No significant effect was found for attendance of religious institution or farmers that occupy some prominent position (e.g. government officials or village chiefs).

5.1.3. Technical Specifications

Seed Type, Seed Source and Age of Trees

In accordance with Richman (2010) we find no evidence of significant effect of usage of hybrid cocoa variants on the TE.

Seed source is one of the new technical variables used in this thesis. It is whether the new seeds are purchased or cocoa is just replanted. And the findings suggest farms using replanting are more efficient than the one buying the seeds.

Other is the age of trees, where farms with older trees are found to be more efficient, this however is not very surprising, since it takes several years for cocoa tree to reach its full yields.

Negative Events

The occurrence of negative events, like black pod, swollen shoot, caterpillars and other insects among other, has negative significant effect on efficiency, as it could be expected. This is also supported by findings of negative effect of difference in number of events on efficiency change. However, no significant effects of dummy variables for the important individual event types were found. This is rather surprising, but it could be caused by relatively lower number of each of the events individually occurring. Only difference in the number of black pod occurrences is found to have significant negative effect on the efficiency change, as reported in the regression (5) in the Table 5.6, thus it can be concluded that black pod attack has negative impact on efficiency. Richman (2010) also found negative impact of events on efficiency, even individually for several types of events, black pod attack among them.

Usage of Inputs

When accounting for village specific effects, farmers using other agricultural equipment (the one not included in capital) were found about 9% more efficient than those who do not use any. This is reasonable finding, because this variable accounts for undervaluation of the capital those farmers use. To see whether there is other effect of usage of other equipment, we would have to include it also to the capital variable, but we do not have data to do so.

Very significant negative effect was found for the usage of insecticides, with scale of some -10%. This could be caused if insecticides would be applied as a response to occurrence of insects, or if insecticides would have some negative effects. To answer this question we would however need to perform more on-field research. With this effect again supported by findings in the efficiency change analysis in the Table 5.6. All this supports the negative estimated effect of insecticides use by Richman (2010).

Richman (2010) also finds fertilizer usage to have positive effect on efficiency; we are, however, unable to support these findings.

Some positive effects of usage of fungicides could be suggested based on our results.

Last the farmers using spray machines are found to be more efficient and farmers using government spray gang to be less efficient. First of these is also supported by TE change analysis results. This supports findings of Richman (2010) who also found positive effect of usage of spraying machine on the TE.

5.1.4. Focus on Farming

This thesis wants to research the a hypothesis whether there is any impact of the focus on farming on efficiency. With the hypothesis being that farmers who focus on (cocoa) farming more should be more efficient. This effect will be researched looking at effect of several proxy variables.

Results of the Tobit regressions performed as shown in the Tables 5.4 and 5.5 suggest that there is a positive effect of focus on farming, more precisely the proxy variable ”main occupation farmer”. Farmers who stated ”farmer” as their main occupation are found more efficient then the others, but this result is significant only in some specifications, however this suggested impact would be very high, over 10%. The effect of difference in stating the main occupation to be farmer on efficiency change is found to be negative, which contradicts the previous findings, however, this change only occurred for about 4% of the observations.

Somewhat stronger results are found for the influence of share of land for cocoa out of the total land on the TE. In the basic specification (1) this effect is found to be some 7.7% and significant on 95% confidence level.

Very surprising is the effect of receiving rental income on efficiency, which is 25% and significant on 99% confidence level. However, only 1.5% of farmers receive rental income, and this farmers use much higher levels of all factors of production. This effect could be result of some more favourable starting position for some of the farmers or to be due to some unobserved difference between the two groups like skill.

Another possible proxy for focus on farming could be number of animals farmer owns. Found to have strongly significant positive effect, this supports the findings for the main occupation dummy variable.

No significant effects are found for buying and selling food dummy variables as for other income dummy.

In conclusion, it rather seems farmers who focus on farming and especially cocoa farming are more technically efficient in cocoa farming, this supports our initial hypothesis.

Other literature does not pay much attention to this issue, but for example Coelli et al. (2002) found negative effect of off-farm income on the TE of Bangladeshi rice farmers, which basically says the same about focus on farming and its effect on efficiency as our findings.

5.1.5. Farm Management

The farms that have a caretaker are on average about 8.5% more efficient. This effect is strongly significant in all the specifications we use.

Usage of paid labour was not found to have any significant effect on the TE. Same is the situation with way annual labourers are paid, paying share of output has no significant effect in comparison with those paying fixed wage.

5.1.6. Investments

Other effects given not so much attention in current literature are the effects of investments. The investments for which we try whether there is any effect on efficiency can be divided into two basic groups. One is investments more explaining focus on farming and possibilities of farmers. In the other are investments connected directly to the cocoa production.

In the first one, investing in non-cocoa land and investing in other businesses is found to have no effect on efficiency. However, investing in buildings dummy is found to have positive and significant effect, it again could be connected to wealth of the farmer, same as for rental income. Quite surprising finding is, however, that farmers investing in children's education are found to be less efficient. Maybe they have fewer resources available for cocoa production or again, they might give less importance to it.

In the second category of investment, farmers who invested in expansion of their farm are found to be less efficient, perhaps because to properly set up the cocoa production takes some time. No effect of investment in replanting dummy on the TE is found.

Table 5.4.: Efficiency determinants Tobit results (1) - (5)

	(1)	(2)	(3)	(4)	(5)
VARIABLES	TE	TE	TE	TE	TE
log(labour)	-0.0684***	-0.0476***	-0.0739***	-0.0982***	-0.115***
	(0.00745)	(0.00694)	(0.00717)	(0.0140)	(0.0143)
log(cocoa area)	0.0381***	0.0367***	0.0339***	0.0201	0.0850***
	(0.0107)	(0.0108)	(0.0103)	(0.0188)	(0.0192)
log(capital +1)	-0.00598	-0.00776**	-0.00789	0.0123	-0.00827
	(0.00554)	(0.00323)	(0.00536)	(0.00948)	(0.0152)
log(inputs +1)	-0.0115	-0.0229***	-0.0185**	-0.00235	0.0189
	(0.00746)	(0.00740)	(0.00722)	(0.0139)	(0.0139)
Age	-0.000536	-0.000962*	0.000657	0.000238	0.000160
	(0.000580)	(0.000581)	(0.000564)	(0.00101)	(0.00104)
Female	-0.0378*	-0.0405*	-0.0178	-0.00539	0.00249
	(0.0224)	(0.0228)	(0.0215)	(0.0407)	(0.0395)
Education	0.0189**	0.0143*	0.0227***	0.0349**	0.0237*
	(0.00777)	(0.00786)	(0.00757)	(0.0136)	(0.0134)
Caretaker	0.0855***	0.0847***	0.0775***	0.153***	0.0717**
	(0.0155)	(0.0157)	(0.0151)	(0.0290)	(0.0322)
Main occupation farmer	0.0415	0.0418	0.0468	0.191***	0.157*
	(0.0385)	(0.0392)	(0.0369)	(0.0716)	(0.0949)
Share land for c.	0.0768**	0.0509	0.0109	0.0646	0.134*
	(0.0378)	(0.0382)	(0.0374)	(0.0659)	(0.0741)
Buys food	-0.00327	0.000547	-0.00178	-0.0476	-0.0265
	(0.0157)	(0.0158)	(0.0152)	(0.0310)	(0.0297)
Sells food	-0.0183	-0.0359	-0.00152	-0.000357	-0.0480
	(0.0237)	(0.0240)	(0.0232)	(0.0433)	(0.0431)
Other income	-0.0167	0.00266	-0.0136	-0.0425	-0.0474
	(0.0150)	(0.0146)	(0.0147)	(0.0304)	(0.0353)
Saves money for coco	0.0440***	0.0491***	0.0487***	0.0581**	0.0552*
	(0.0148)	(0.0148)	(0.0142)	(0.0279)	(0.0291)
Borrows money for coco	-0.0312*	-0.0329**	-0.0253	-0.00452	-0.00734
	(0.0162)	(0.0165)	(0.0155)	(0.0302)	(0.0358)
Used - other agri. equip.	0.0443	-0.00210	0.0931***	0.0394	-0.113
	(0.0343)	(0.0337)	(0.0337)	(0.0515)	(0.124)
Used - fertilizer	-0.0108	0.00168	-0.0228	0.0277	0.00902
	(0.0178)	(0.0180)	(0.0175)	(0.0342)	(0.0342)
Used - insecticides	-0.113***	-0.0961***	-0.104***	-0.138***	-0.0991*
	(0.0256)	(0.0259)	(0.0247)	(0.0526)	(0.0516)
Used - fungicides	0.0175	0.0308**	0.0149	-0.00653	-0.0108
	(0.0156)	(0.0156)	(0.0150)	(0.0297)	(0.0291)

	(1)	(2)	(3)	(4)	(5)
Used - sp. m. non-gov.	0.0287	0.0509***	0.0210	-0.0316	-0.0596
	(0.0188)	(0.0188)	(0.0181)	(0.0353)	(0.0419)
Used - gov. spray gang	-0.0549***	-0.0394**	-0.0158	-0.0384	-0.0667**
	(0.0181)	(0.0183)	(0.0188)	(0.0327)	(0.0313)
Used - other inputs	-0.00951	-0.00672	-0.0281	0.143	-0.165
	(0.0669)	(0.0680)	(0.0644)	(0.182)	(0.145)
Proportion land own	0.0563**	0.0516*	0.0259	-0.0138	0.0312
	(0.0271)	(0.0275)	(0.0263)	(0.0504)	(0.0551)
Share hybrid	0.00832	0.000697	0.00254	-0.0139	0.0540*
	(0.0157)	(0.0159)	(0.0158)	(0.0290)	(0.0289)
Event nu.	-0.0215**	-0.0115	-0.0163*	-0.0172	0.0341
	(0.00908)	(0.00892)	(0.00877)	(0.0180)	(0.0208)
Paid labour usage	0.00531	0.0190	7.12e-06	0.00504	0.0440
	(0.0214)	(0.0216)	(0.0207)	(0.0380)	(0.0374)
2006	-0.106***		-0.113***		
	(0.0232)		(0.0228)		
2008	-0.130***		-0.133***		
	(0.0265)		(0.0259)		
2010	-0.153***		-0.136***	-0.0932*	
	(0.0257)		(0.0255)	(0.0545)	
Village and region dummies			YES***		
Animals				0.00187***	
				(0.000429)	
Invest - landother					-0.0198
					(0.0697)
Invest - buildings					0.0756**
					(0.0384)
Invest - childeduc					-0.0733**
					(0.0313)
Constant	0.734***	0.570***	0.652***	0.578***	0.404**
	(0.0831)	(0.0807)	(0.0905)	(0.141)	(0.163)
sigma_u	0.124***	0.127***	0.0922***	0.0920**	4.79e-10
	(0.0107)	(0.0110)	(0.0119)	(0.0433)	(39.80)
sigma_e	0.216***	0.219***	0.212***	0.230***	0.238***
	(0.00612)	(0.00627)	(0.00597)	(0.0185)	(0.00909)
Observations	1,373	1,373	1,367	425	384
Number of farmers	711	711	707	338	384

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.5.: Efficiency determinants Tobit results (6) - (10)

	(6)	(7)	(8)	(9)	(10)
VARIABLES	TE	TE	TE	TE	TE
log(labour)	-0.0749***	-0.0549***	-0.0771***	-0.0425***	-0.0678***
	(0.00838)	(0.00915)	(0.00829)	(0.0116)	(0.00747)
log(cocoa area)	0.0377***	0.0239*	0.0374***	0.0703***	0.0416***
	(0.0117)	(0.0128)	(0.0117)	(0.0151)	(0.0107)
log(capital +1)	-0.000524	-0.00299	-0.00152	-0.0217***	-0.00477
	(0.00614)	(0.00639)	(0.00612)	(0.00777)	(0.00550)
log(inputs +1)	-0.00854	-0.0139	-0.00975	-0.00453	-0.0269***
	(0.00803)	(0.00952)	(0.00808)	(0.0107)	(0.00459)
Age	-3.99e-05	-0.000648	-0.000256	-0.000353	-0.000481
	(0.000631)	(0.000670)	(0.000638)	(0.000796)	(0.000585)
Female	-0.0297	-0.0453*	-0.0272	-0.0337	-0.0431*
	(0.0246)	(0.0255)	(0.0246)	(0.0321)	(0.0227)
Education	0.0246***	0.0160*	0.0234***	0.0216**	0.0183**
	(0.00853)	(0.00885)	(0.00848)	(0.0107)	(0.00786)
Caretaker	0.0859***	0.0891***	0.0840***	0.0634***	0.0850***
	(0.0170)	(0.0182)	(0.0170)	(0.0214)	(0.0156)
Main occupation farmer	0.0403	-0.00104	0.0413	0.0882*	0.0528
	(0.0426)	(0.0418)	(0.0425)	(0.0505)	(0.0388)
Share land for c.	0.0551	0.134***	0.0499	0.158***	0.0881**
	(0.0432)	(0.0436)	(0.0431)	(0.0498)	(0.0379)
Buys food	-0.0155	0.0128	-0.0159	0.00662	-0.000867
	(0.0181)	(0.0200)	(0.0179)	(0.0227)	(0.0158)
Sells food	-0.00156	-0.0251	-0.00334	-0.0257	-0.00940
	(0.0257)	(0.0269)	(0.0257)	(0.0289)	(0.0238)
Other income	-0.00953	-0.0340*	-0.0126	0.00316	-0.0211
	(0.0171)	(0.0195)	(0.0170)	(0.0212)	(0.0151)
Saves money for coco	0.0455***	0.0426**	0.0485***	0.0410**	0.0398***
	(0.0169)	(0.0172)	(0.0168)	(0.0201)	(0.0148)
Borrows money for coco	-0.0318*	-0.0268	-0.0316*	0.00225	-0.0334**
	(0.0185)	(0.0185)	(0.0184)	(0.0215)	(0.0163)
Used - other agri. equip.	0.0507	0.0484	0.0502	0.129**	0.0555
	(0.0342)	(0.0411)	(0.0341)	(0.0520)	(0.0343)
Used - fertilizer	-0.0133	-0.0216	-0.0134	-0.0161	
	(0.0204)	(0.0212)	(0.0203)	(0.0239)	
Used - insecticides	-0.111***	-0.0839***	-0.111***	-0.123***	
	(0.0264)	(0.0318)	(0.0264)	(0.0372)	
Used - fungicides	-0.00388	0.0309*	-0.00217	0.0145	
	(0.0179)	(0.0184)	(0.0179)	(0.0216)	
Used - spray m. non-gov.	0.0228	0.0390*	0.0289	0.0416*	
	(0.0197)	(0.0222)	(0.0198)	(0.0243)	

	(6)	(7)	(8)	(9)	(10)
Used - gov. spray gang	-0.0681***	-0.0440**	-0.0693***	-0.0248	
	(0.0193)	(0.0223)	(0.0193)	(0.0251)	
Used - other inputs	-0.0137	0.0700	-0.0182	-0.0339	
	(0.0663)	(0.114)	(0.0661)	(0.122)	
Proportion land own	0.0728**	0.0243	0.0745**	0.0875**	0.0634**
	(0.0308)	(0.0311)	(0.0308)	(0.0398)	(0.0272)
Share hybrid	0.00634	0.00567	0.0204	-0.0127	0.00801
	(0.0175)	(0.0183)	(0.0182)	(0.0221)	(0.0159)
Event nu.	-0.0177	-0.0360***	-0.0185*	-0.0122	-0.0192**
	(0.0111)	(0.0105)	(0.0110)	(0.0126)	(0.00912)
Paid labour usage	0.00295	-0.0178	0.00751	0.0298	0.00641
	(0.0221)	(0.0278)	(0.0221)	(0.0322)	(0.0215)
2006		-0.105***		-0.0433	-0.0951***
		(0.0242)		(0.0300)	(0.0224)
2008	-0.0113	-0.0947***	-0.0156	-0.169***	-0.124***
	(0.0336)	(0.0318)	(0.0335)	(0.0367)	(0.0262)
2010	-0.0522***	-0.119***	-0.0536***	-0.0545	-0.138***
	(0.0199)	(0.0387)	(0.0197)	(0.0354)	(0.0251)
Invest - expansion	-0.0423*				
	(0.0218)				
Invest - coco. replanting	0.00839				
	(0.0213)				
Invest - other business	0.0224				
	(0.0434)				
Rental income		0.236***			
		(0.0692)			
Age of trees			0.00345***		
			(0.00128)		
Seeds - replacing share				0.0401*	
				(0.0210)	
Constant	0.614***	0.715***	0.588***	0.227*	0.636***
	(0.0884)	(0.0975)	(0.0888)	(0.123)	(0.0812)
sigma_u	0.124***	0.116***	0.125***	0.119***	0.128***
	(0.0120)	(0.0162)	(0.0120)	(0.0200)	(0.0106)
sigma_e	0.211***	0.201***	0.211***	0.168***	0.217***
	(0.00719)	(0.00895)	(0.00718)	(0.0124)	(0.00614)
Observations	1,080	876	1,079	510	1,376
Number of farmers	649	572	649	379	712

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.2. Determinants of Efficiency Change

Current literature does not research the issue of farmers' efficiency change and its determinants much. Some studies are researching the marginal effects of the factors they use in determining TE scores, examples of such studies are Dhungana et al. (2004) or Haji (2007). However, no study found was trying to look on some factors whose static levels would influence efficiency scores.

To investigate this issue we again perform total of fourteen regressions, this time OLS as explained in the methodology part of this thesis. First regressions (1) includes social variables that are suspected to possibly have some effect in determining TE change, next it also includes differences in of other variables used in the model created to explain TE. The second (2) and the third (3) regressions differ from the first one in non-inclusion of year dummy variable and in extra inclusion of village and regional dummy variables respectively. Regressions (3) to (5) include extra variables for investments and for difference in negative events. The results of the first five regressions are reported in Table 5.6. Other eight regressions, found separately no significant effect for differences in way annual labourers are paid, rental income, receiving something for free and or found no significant effect for values of being a prominent, attendance of religious institution, age of trees, seeds replacing share, investments in other land, children education and buildings. These regressions are not reported. Lastly, significant positive effect was found for difference in number of animals, however, this regression only used very few observations and the differences in numbers of animals are huge and unexpected, since the variable is constructed in quite difficult way. This regression is also not being reported.

Marginal effects are already individual commented upon in the part about determinants of efficiency, and are used to support the findings there, this section will only comment on the determinants of efficiency change.

5.2.1. Farmer Characteristics and Efficiency Change

Age

The age if found to have significant positive effect on efficiency change. This suggest that older farmers tend to get more efficient, possibly since they gained more experience they can use to improve their production process. This explanation however is not very strong, since older farmers are not found to be on average more efficient. To properly investigate this issue panel from longer time period would be needed.

Gander

It is found that female farmers are getting significantly more efficient in time in comparison with the male farmers. This effect is found to be significant even on at the 99% confidence level in the specification (1). Moreover this effect is of a very large scale of 7.8%. This finding is in a way in opposition with the finding that female farmers are less efficient.

When taking into account the average differences in factors of production used by male and female farmers do not differ greatly, it seems reasonable to accept these findings.

To conclude that it seems that even though, the females are less efficient (possibly due to facing unfavourable conditions as discussed previously in this thesis) they might be better in learning and improving in time. And thus, female empowerment could help cocoa production in Ghana.

Saving

We found that farmers saving for cocoa production on average get more efficient in time then the rest of the farmers. This effect even is significant at 99% confidence level in four out of five regressions reported. Moreover the suggested effect is higher 6%, which again is already quite an impact. Similar explanation as for the effect of savings on the level of TE can be used. Maybe, the farmers that are able to save are the more capable ones that also tend to get more efficient over time. Or possibly, the savings give the farmers the resources to operate more efficiently and to improve in time.

Other

No significant effects were found of other social variables on efficiency change. These include being a prominent, attending a religious institution, level of education and borrowing money for cocoa production.

5.2.2. Investments and Efficiency Change

We decided to put investments as a levels and not differences in our analysis because the result of such investments would not be yield results immediately anyway.

Cocoa replanting

The farmers investing in replanting are found to get more efficient in time then those who do not invest as can be seen in regression (4) in the Table 5.6. Since the dummy for investment in replanting was obtained by asking the farmers question whether they invested in replanting in last few years, this effect could also be caused simply by the new trees finally coming to yield its full production and thus making the farm more efficient then in the previous period.

Other

No other investments were found to have significant effect on efficiency change. These include investments in expansion, other business, buildings, other land and children education.

Table 5.6.: Efficiency change determinants OLS results

	(1)	(2)	(3)	(4)	(5)
VARIABLES	TE change	TE change	TE change	TE change	TE change
VALUES					
Age	0.00138*	0.00170**	0.00121	0.00149*	0.00139*
	(0.000824)	(0.000816)	(0.000917)	(0.000825)	(0.000826)
Female	0.0780**	0.0790**	0.0757**	0.0805**	0.0758**
	(0.0341)	(0.0343)	(0.0369)	(0.0342)	(0.0342)
Education - level	0.0112	0.0120	0.0121	0.0114	0.0109
	(0.0113)	(0.0114)	(0.0128)	(0.0115)	(0.0114)
Share land for c.	0.0294	0.0487	0.0408	0.0240	0.0176
	(0.0758)	(0.0763)	(0.0866)	(0.0759)	(0.0760)
Saves money for coco	0.0631***	0.0593**	0.0623**	0.0642***	0.0638***
	(0.0240)	(0.0242)	(0.0249)	(0.0242)	(0.0242)
Borrows money for coco	0.00850	0.00284	-0.00571	-0.000423	0.00589
	(0.0336)	(0.0338)	(0.0349)	(0.0337)	(0.0337)
DIFFERENCES					
log(labour)	-0.0665***	-0.0615***	-0.0683***	-0.0681***	-0.0685***
	(0.00898)	(0.00893)	(0.00913)	(0.00905)	(0.00914)
log(cocoa area)	-0.0129	-0.0155	-0.0107	-0.0124	-0.0163
	(0.0212)	(0.0213)	(0.0217)	(0.0212)	(0.0212)
log(capital +1)	-0.0269***	-0.0118***	-0.0259***	-0.0273***	-0.0253***
	(0.00577)	(0.00340)	(0.00600)	(0.00579)	(0.00584)
log(inputs +1)	-0.0388***	-0.0405***	-0.0361***	-0.0381***	-0.0386***
	(0.00893)	(0.00899)	(0.00907)	(0.00892)	(0.00894)
Caretaker	0.0154	0.00629	0.0182	0.0168	0.0149
	(0.0198)	(0.0198)	(0.0201)	(0.0199)	(0.0200)
Main occupation farmer	-0.0876*	-0.0788	-0.0952*	-0.0879*	-0.0821
	(0.0518)	(0.0522)	(0.0527)	(0.0517)	(0.0520)
Share land for c.	-0.0124	-0.0281	-0.0250	-0.0172	-0.0146
	(0.0606)	(0.0608)	(0.0656)	(0.0607)	(0.0609)
Other income	0.0274	0.0217	0.0272	0.0286	0.0259
	(0.0180)	(0.0180)	(0.0185)	(0.0181)	(0.0181)
Borrows money for coco	-0.0559**	-0.0555**	-0.0504**	-0.0500**	-0.0511**
	(0.0250)	(0.0252)	(0.0254)	(0.0251)	(0.0252)
Used - other agri. equip.	0.0232	-0.00364	0.0458	0.0239	0.0231
	(0.0373)	(0.0369)	(0.0379)	(0.0373)	(0.0377)
Used - fertilizer	-0.00455	-0.00225	-0.00985	-0.00503	-0.00607
	(0.0219)	(0.0221)	(0.0226)	(0.0219)	(0.0220)
Used - insecticides	-0.0554*	-0.0552*	-0.0663**	-0.0576**	-0.0558*
	(0.0288)	(0.0290)	(0.0296)	(0.0288)	(0.0289)
Used - fungicides	0.0269	0.0380**	0.0270	0.0244	0.0300
	(0.0189)	(0.0188)	(0.0195)	(0.0190)	(0.0190)

	(1)	(2)	(3)	(4)	(5)
Used - spray m. non-gov.	0.0540**	0.0583***	0.0550**	0.0525**	0.0501**
	(0.0211)	(0.0210)	(0.0214)	(0.0211)	(0.0212)
Used - gov. spray gang	-0.0103	-0.00898	-0.00496	-0.0123	-0.00886
	(0.0231)	(0.0233)	(0.0238)	(0.0231)	(0.0232)
Used - other inputs	0.0629	0.0621	0.0515	0.0643	0.0605
	(0.0644)	(0.0650)	(0.0651)	(0.0643)	(0.0645)
Share hybrid	-0.0363	-0.0285	-0.0298	-0.0375	-0.0371
	(0.0258)	(0.0259)	(0.0265)	(0.0259)	(0.0264)
Event nu.	-0.0207*	-0.0224**	-0.0232**	-0.0213*	
	(0.0111)	(0.0106)	(0.0113)	(0.0111)	
Paid labour usage	-0.00414	-0.00211	-0.0149	0.000288	-0.00449
	(0.0252)	(0.0254)	(0.0262)	(0.0253)	(0.0253)
Proportion land own	0.0676*	0.0623	0.0637	0.0653	0.0625
	(0.0397)	(0.0400)	(0.0403)	(0.0398)	(0.0398)
Difference over 2 periods	-0.0234	0.0262	-0.0150	-0.0248	-0.0300
	(0.0354)	(0.0330)	(0.0370)	(0.0356)	(0.0355)
Difference over 3 periods	-0.301***	-0.218***	-0.287***	-0.303***	-0.314***
	(0.0833)	(0.0807)	(0.0864)	(0.0832)	(0.0836)
2008	-0.0682		-0.0564	-0.0729	-0.0631
	(0.0447)		(0.0482)	(0.0446)	(0.0448)
2010	0.112***		0.108***	0.107***	0.117***
	(0.0355)		(0.0380)	(0.0357)	(0.0363)
Village and regional dummies			YES		
Invest - expansion				-0.0371	
				(0.0302)	
Invest - coco. replanting				0.0684**	
				(0.0309)	
Invest - other business				-0.0332	
				(0.0611)	
Event - black pod - diff.					-0.0351*
					(0.0195)
Event - swollen shoot - diff.					-0.0153
					(0.0364)
Event - caterpillar - diff.					0.0456
					(0.0320)
Event - other insects- diff.					-0.0388
					(0.0313)
Event - other - diff.					-0.0190
					(0.0243)
Constant	-0.186**	-0.212**	-0.337**	-0.186**	-0.182**
	(0.0891)	(0.0885)	(0.134)	(0.0895)	(0.0892)
Observations	655	655	653	653	655
R-squared	0.254	0.239	0.293	0.263	0.260

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

5.3. Scale Efficiency

5.3.1. Scores

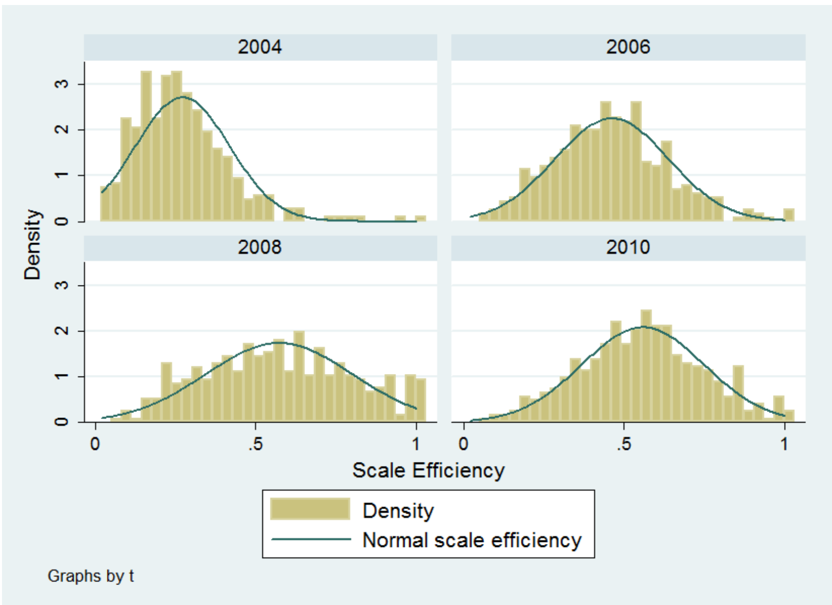
The mean scale efficiency found in our sample is 47.1%. Moreover the SE was increasing in time. Rising from 27.2% in 2004 to 55.9% in 2010 as can be seen in the Table 5.7. The mean scale efficiency in 2008 was even higher, but there is still the problem with much restricted variable capital, so SE scores in 2008 are less comparable with other years, as can be also seen on much higher number of peers. Just to remind, observations from 2008 were not allowed to serve as peers for the other years. Moreover the standard deviation in scores from 2008 is higher then in other years.

Table 5.7.: Summary statistics of the scale efficiency scores

Scale efficiency	n	mean	sd	min	max	nu. of peers
2004	339	0.272	0.147	0.020	1	1
2006	364	0.462	0.177	0.070	1	3
2008	370	0.572	0.230	0.077	1	11
2010	389	0.559	0.192	0.068	1	3
total	1462	0.471	0.223	0.020	1	18

The distribution of scale efficiency in different years quite follows a normal distribution, as is shown in the Figure 5.4.

Figure 5.4.: Scale efficiency distribution in years



The scale efficiency scores we found are much lower than the ones found by other studies on farmers' efficiency. Eyitayo et al. (2011) finds scale efficiency of about 70% for Nigerian cocoa farmers, while the other studies of efficiency of cocoa farmers do not investigate the issue of scale efficiency (because they use SFA). Other studies also found lower SE scores, for example: Binam et al. (2003) finds SE of 77% for coffee farmers from Côte dIvoire, Dhungana et al. (2004) finds SE of 93% for Nepalese rice farmers and Coelli et al. (2002) also finds scale efficiency of about 93% for rice farmers, in this case from Bangladesh.

5.3.2. Determinants

We only looked at the determinants of scale efficiency in a way to investigate effect of the variables used for analyzing determinants of the TE while controlling for factors of production, which are suspected to effect the SE in complex way, we will include all linear, quadratic and logarithmic values for factors of production.

We again performed multiple regressions in same specification as when researching determinants of TE only extra including the variables of the factors of production and their squares. Reported are only the three most basic regressions, since the other found no significant effects. Second and third regression differ from the first one in the way that second (2) does not include time dummy variables and third (3) includes extra village and regional dummy variables. The results of regressions (1) to (3) are reported in the Table 5.8.

Age

Significant positive effect of age of farmers on scale effiiciency was found. But this effect is so small that it has literally no real life impact, since 10 year older farmer would not be even 1% more scale efficient on average.

Gender

It was found that female farmers are significantly less scale efficient than males. The real life impact is considerable, about 3% on average. And it is found to be significant with at last 90% confidence level in all three specifications. This further supports the hypothesis that female farmers could face less favourable conditions and be disadvantaged.

Education

Education is found to have significant positive effect on the scale efficiency. This further supports the hypothesis that more educated farmers are the more efficient ones, and not just technically but also scale efficient. However, the effect of education on scale efficiency is not very high, extra level of education increasing on average the SE by just something above 0.7%.

Usage of Inputs

The results suggest that the farmers that used assistance of the government spray gang are less scale efficient. This could be again due to some unobserved effect, like skill of farmers that might differ for the two categories. Usage of non-government supplied spraying machine exhibits negative effect, but significant only for specification without time dummy variables, since the level of usage is a bit higher in 2004 and since the SE is much lower in 2004, it just makes it seem significant when no time dummies are used, so it is no strong finding. Same it is for of usage of other agricultural equipment that is higher in 2008, the year with the highest SE scores.

Focus on Farming

Farmers that buy food are found to be more scale efficient. With this effect being significant with at last 90% confidence level in all three specifications. This could be because, the farmers who buy food are more focused on cocoa farming, and focusing on cocoa farming would allow them to be more scale efficient. However this is not much supported by other findings, since the effect of share of land for cocoa is only found significant in the specification without time dummy variables, but there its significance is likely to be caused by slightly lower levels of the share in 2004, similarly as for the investment variables.

Moreover, to even more properly investigate all the above stated effects, we would have to conduct deeper analysis, that would also for example looked on whether the farmers operate in the increasing or decreasing returns to scale spectrum.

Table 5.8.: Scale efficiency determinants Tobit results

	(1)	(2)	(3)
VARIABLES	SE	SE	SE
Labour	-0.000121***	-0.000146***	-0.000126***
	(2.19e-05)	(2.41e-05)	(2.19e-05)
(labour)^2	2.12e-08***	2.41e-08***	2.20e-08***
	(4.06e-09)	(4.48e-09)	(4.07e-09)
log(labour)	-0.102***	-0.114***	-0.101***
	(0.00588)	(0.00651)	(0.00589)
Cocoa area	-0.0103***	-0.0106***	-0.00993***
	(0.000921)	(0.00102)	(0.000934)
(cocoa area)^2	3.47e-05***	3.54e-05***	3.20e-05***
	(5.18e-06)	(5.73e-06)	(5.24e-06)
log(cocoa area)	0.210***	0.217***	0.208***
	(0.0114)	(0.0126)	(0.0117)
Capital	0.000106***	0.000360***	8.68e-05***
	(3.18e-05)	(3.12e-05)	(3.20e-05)
(capital)^2	-2.91e-08**	-1.06e-07***	-2.38e-08**
	(1.20e-08)	(1.23e-08)	(1.20e-08)
log(capital +1)	0.0272***	-0.0161***	0.0281***
	(0.00390)	(0.00256)	(0.00392)
Inputs	-8.51e-07	3.09e-06	-4.72e-07
	(1.01e-05)	(1.12e-05)	(1.02e-05)
(inputs)^2	-5.65e-10	-3.74e-10	-5.67e-10
	(6.21e-10)	(6.86e-10)	(6.21e-10)
log(inputs +1)	-0.0405***	-0.0333***	-0.0433***
	(0.00446)	(0.00489)	(0.00450)
Age	0.000395	0.000661**	0.000491*
	(0.000255)	(0.000278)	(0.000270)
Gender - female	-0.0292***	-0.0391***	-0.0239**
	(0.00991)	(0.0109)	(0.0103)
Education	0.00723**	0.0112***	0.00773**
	(0.00347)	(0.00382)	(0.00367)
Caretaker	0.00417	0.00866	0.00559
	(0.00742)	(0.00818)	(0.00763)
Main occupation farmer	0.0104	0.00524	0.00906
	(0.0184)	(0.0203)	(0.0186)
Share land for cocoa	0.0279	0.0398**	0.0296
	(0.0180)	(0.0198)	(0.0190)
Buys food	0.0189**	0.0229***	0.0173**
	(0.00781)	(0.00853)	(0.00788)
Sells food	-0.00385	0.0144	-0.00222
	(0.0117)	(0.0128)	(0.0120)

	(1)	(2)	(3)
Other income	-0.0100	-0.0174**	-0.0136*
	(0.00731)	(0.00777)	(0.00750)
Saves money for coco	0.00901	0.0157**	0.00848
	(0.00719)	(0.00787)	(0.00726)
Borrows money for coco	0.000463	0.00439	0.00227
	(0.00784)	(0.00867)	(0.00788)
Used - other agri. equip.	-0.0104	0.0690***	-0.00380
	(0.0170)	(0.0181)	(0.0174)
Used - fertilizer	0.0224**	0.0123	0.0268***
	(0.00872)	(0.00961)	(0.00902)
Used - insecticides	0.0187	0.0138	0.0231*
	(0.0132)	(0.0145)	(0.0132)
Used - fungicides	-0.000148	-0.0120	-0.000905
	(0.00767)	(0.00840)	(0.00773)
Used - spray m. non-gov.	-0.0107	-0.0278***	-0.0103
	(0.00931)	(0.0102)	(0.00939)
Used - gov. spray gang	-0.0150*	-0.0305***	-0.00665
	(0.00878)	(0.00964)	(0.00966)
Used - other inputs	0.0202	0.0418	0.0163
	(0.0336)	(0.0372)	(0.0335)
Proportion land own	-0.00479	0.00665	0.00268
	(0.0129)	(0.0142)	(0.0133)
Share hybrid	-0.0120*	-0.0122	-0.0114
	(0.00724)	(0.00796)	(0.00781)
Event nu.	-0.00183	-0.00326	-0.00204
	(0.00449)	(0.00480)	(0.00454)
Paid labour usage	0.00165	-0.0122	0.00281
	(0.0105)	(0.0116)	(0.0107)
2006	0.0415***		0.0415***
	(0.0119)		(0.0121)
2008	0.262***		0.265***
	(0.0151)		(0.0153)
2010	0.0703***		0.0745***
	(0.0129)		(0.0134)
Village dummy			YES
Constant	0.588***	0.832***	0.580***
	(0.0449)	(0.0473)	(0.0501)
sigma_u		0	0
	(0.0151)	(0.0138)	(0.0119)
sigma_e	0.119***	0.131***	0.117***
	(0.00228)	(0.00253)	(0.00226)
Observations	1,373	1,373	1,367
Number of s0far	711	711	707

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter 6
Conclusion

This thesis investigated technical efficiency, its change and scale efficiency and the determinants of these on a unbalanced panel of more than 700 Ghanaian cocoa farmers collected between 2004 and 2010 in several regions of Ghana.

To find the efficiency scores we used output-oriented variable returns to scale data envelopment analysis model. The mean technical efficiency found in our sample was very low, just 34.5%. Moreover, efficiency was found to change quite a lot in time, the mean of absolute value of efficiency change in our sample was 21%. However, the DEA model has its drawbacks and the mean TE scores are likely to be downward biased.

To investigate the determinants of efficiency we used random effects Tobit regression on the TE scores. Several variables were found to have significant effect on technical efficiency. Among the ones with positive effect are farms size, farmers' education, saving money for cocoa production, ownership of land, usage of spray machine, proxies for focus on farming, receiving a rental income, having a caretaker, investing in buildings, age of trees and replacing as seed source. Among the ones with negative effect are gender - female, borrowing money for cocoa production, usage of government spray gang and insecticides, occurrence of negative events and investing in expansion.

The major contribution of this thesis is in investigating determinants of efficiency change; apart differences in some of the above stated variables, some variables in their levels were found to have a significant effect on efficiency change. Older farmers were found to get more efficient in time, so were female farmers and farmers saving for cocoa production. No significant determinants of efficiency change with negative effect were found.

One policy implication of these findings could be to work on improving of the position of female farmers. Female farmers were found to be less efficient then male farmers, possibly due to facing unfavourable conditions. However, they are more able to learn and improve over time significantly more than men.

Research conducted by this thesis could possibly be extended with more on-level research to deeper investigate some of the effects found, especially effects like the negative impact of credit on efficiency. Furthermore, data could be collected with focus on conducting a efficiency study and making sure there is not so many missing and clearly wrong values. Moreover, the whole analysis could be conducted using stochastic frontier analysis to compare with our results to see whether the effects would be still the same.

Bibliography

Amos, T. T., 2007. An analysis of productivity and technical efficiency of smallholder cocoa farmers in nigeria. Journal of Social Sciences 15 (2), 127–133.

Aneani, F., Anchirinah, V. M., Asamoah, M., Owusu-Ansah, F., 2011. Analysis of economic efficiency in cocoa production in ghana. African Journal of Food, Agriculture, Nutrition and Development 11 (1).

Binam, J. N., Gockowski, J., Nkamleu, G. B., 2008. Technical efficiency and productivity potential of cocoa farmers in west african countries. The Developing Economies 46 (3), 242–263.

Binam, J. N., Sylla, K., Diarra, I., Nyambi, G., 2003. Factors affecting technical efficiency among coffee farmers in cote d'Ivoire: evidence from the centre west region. African Development Review 15 (1), 66–76.

Biørn, E., Hagen, T. P., Iversen, T., Magnussen, J., Nov. 2003. The effect of activity-based financing on hospital efficiency: A panel data analysis of DEA efficiency scores 1992–2000. Health Care Management Science 6 (4), 271–283.

Bravo-Ureta, B. E., Evenson, R. E., 1994. Efficiency in agricultural production: the case of peasant farmers in eastern paraguay. Agricultural economics 10 (1), 27–37.

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

Charnes, A., Cooper, W. W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European journal of operational research 2 (6), 429–444.

Coelli, T., Rahman, S., Thirtle, C., 2002. Technical, allocative, cost and scale efficiencies in bangladesh rice cultivation: A non-parametric approach. Journal of Agricultural Economics 53 (3), 607–626.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., Battese, G. E., Dec. 2005. An Introduction to Efficiency and Productivity Analysis. Springer.

CSAE, COCOBOD, 2006. Guide to questionnaires and supporting materials ghana cocoa farmers survey 2006.
URL http://www.andrewzeitlin.com/data/gcfs/GuidetoStructureofGCFS2006.doc?attredirects=0

Danso-Abbeam, G., Aidoo, R., Agyemang, K. O., Ohene-Yankyera, K., 2012. Technical efficiency in ghana's cocoa industry: Evidence from bibiani-anhwiaso-bekwai district. Journal of Development and Agricultural Economics 4 (10), 287–294.

Dhungana, B. R., Nuthall, P. L., Nartea, G. V., 2004. Measuring the economic inefficiency of nepalese rice farms using data envelopment analysis. Australian Journal of Agricultural and Resource Economics 48 (2), 347–369.

Dzokoto, V. A. A., Mensah, E. C., 2010. Making sense of a new currency: An exploration of ghanaian adaptation to the new ghana cedi. Journal of Applied Business & Economics 10 (5).

Eyitayo, O. A., Chris, O., Ejiola, M. T., Enitan, F. T., 2011. Technical efficiency of cocoa farms in cross river state, nigeria. African Journal of Agricultural Research 6 (22), 5080–5086.

Ferrier, G. D., Lovell, C. K., 1990. Measuring cost efficiency in banking: econometric and linear programming evidence. Journal of econometrics 46 (1), 229–245.

Ghana Cocoa Board, 2014. Ghana cocoa board statistics.
URL https://www.cocobod.gh/weakly_purchase.php

Ghana Statistical Service, 2014. CountrySTAT - ghana.
URL http://www.countrystat.org/home.aspx?c=GHA

Haji, J., 2007. Production efficiency of smallholders' vegetable-dominated mixed farming system in eastern ethiopia: A non-parametric approach. Journal of African Economies 16 (1), 1–27.

Helfand, S. M., Levine, E. S., 2004. Farm size and the determinants of productive efficiency in the brazilian center-west. Agricultural Economics 31 (2-3), 241–249.

International Cocoa Organization, 2014. International cocoa organization statistics - production.
URL http://www.icco.org/about-us/international-cocoa-agreements/cat_view/30-related-documents/46-statistics-production.html

McDonald, J., 2009. Using least squares and tobit in second stage DEA efficiency analyses. European Journal of Operational Research 197 (2), 792–798.

Onumah, J. A., Al-Hassan, R. M., Onumah, E. E., 2013. Productivity and technical efficiency of cocoa production in eastern ghana. Journal of Economics and Sustainable Development 4 (4), 106–117.

OSDEA, 2014. Open source DEA.
URL http://www.opensourcedea.org/index.php?title=Open_Source_DEA&oldid=371

Richman, D., 2010. What drives efficiency on the ghanian cocoa farm?
URL http://www.csae.ox.ac.uk/conferences/2010-edia/papers/498-Dzene.pdf

Simar, L., Wilson, P. W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of econometrics 136 (1), 31–64.

Simar, L., Wilson, P. W., Oct. 2011. Two-stage DEA: caveat emptor. Journal of Productivity Analysis 36 (2), 205–218.

Teal, F., Vigneri, M., 2004. Production changes in Ghana cocoa farming households under market reforms. Centre for the Study of African Economies, University of Oxford.
URL http://r4d.dfid.gov.uk/PDF/Outputs/Mis_SPC/R8078a.pdf

World Bank, 2013a. World bank database - employment in agriculture (% of total employment).
URL http://data.worldbank.org/indicator/SL.AGR.EMPL.ZS

World Bank, 2013b. World bank database - inflation, consumer prices (annual %).
URL http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG/

World Bank, 2014. World bank database - agriculture, value added (% of GDP).
URL http://data.worldbank.org/indicator/NV.AGR.TOTL.ZS

Zeitlin, A., 2005. Market structure and productivity growth in ghanaian cocoa production. Centre for the Study of African Economies, University of Oxford.
URL http://www.csae.ox.ac.uk/conferences/2006-EOI-RPI/papers/csae/zeitlin.pdf

Zeitlin, A., Oct. 2008. GCFS 2008 sampling. CSAE internal document.

Appendix A

Tables and Figures

Table A.1.: Mean capital prices and inflation

	2004	2005	2006	2007	2008	2009	2010
Inflation	12.62	15.12	10.92	10.73	16.52	19.25	10.71
Sprey machines	169.64		235.78		245.95		277.11
Pruners	6.12		12.28		20.33		
Cocoa harvesters	1.02		1.53		3.32		
Cutlasses	2.77		3.29		5.51		
Falling axes	3.74		5.29		8.34		

Table A.2.: Mean inputs prices

	2004	2006	2008	2010
Fertilizer (50 kg bags)	22.48	23.94	18.72	19.31
Insecticide (Litres)	9.66	13.37	19.31	15.50
Fungnicide (Sachets)	1.85	1.38	1.51	2.13
Gov. Spray gang (Visit)	2.32	2.72	1.94	3.39
Spray machine rent (Unit)	7.16	8.84	4.92	13.40

Table A.3.: Summary statistics

	count	mean	sd	min	max
Cocoa	2071	1715.2610	2231.4830	0	25000
Cocoa Area	2005	17.5431	19.1351	1	200
Yield / ha	1991	294.7878	265.1456	0	2500
Capital	1885	212.1325	389.5685	0	3498.43
Labour	1813	342.4619	491.7998	1	7180
Inputs	1784	321.6505	852.0679	0	19369.68
Share land for cocoa	2018	0.8575	0.2135	0	1
Plots area total	2041	20.3504	21.9044	0	293
Age	2138	54.7806	14.9060	19	116
Gender - female	1999	0.1621	0.3686	0	1
Education level	1996	1.4424	1.0909	0	5
Prominent	1214	0.2562	0.4367	0	1
Main occupation farmer	1999	0.9630	0.1889	0	1
Church and mosque	622	0.8714	0.3350	0	1
Save money for cocoa	2031	0.3550	0.4786	0	1
Bank account	2033	0.5224	0.4996	0	1
Borrowed for cocoa	2033	0.2223	0.4159	0	1
Proportion land own	2014	0.9133	0.2545	0	1
Annual paid share	643	0.8491	0.3582	0	1
Other income	2030	0.4704	0.4992	0	1
Rental income	1394	0.0143	0.1190	0	1
Other agricultural equip.	2034	0.0531	0.2243	0	1
Animals	566	29.6714	38.2200	0	555
Invest expansion	1520	0.1539	0.3610	0	1
Invest replanting	1520	0.1559	0.3629	0	1
Invest other business	1518	0.0296	0.1697	0	1
Invest land other	469	0.0469	0.2117	0	1
Invest buildings	469	0.1365	0.3436	0	1
Invest children education	900	0.6733	0.4693	0	1
Free fertilizer	808	0.0173	0.1306	0	1
Free insecticides	1634	0.0410	0.1984	0	1
Free fungicides	705	0.2014	0.4013	0	1
Free spray machine	1092	0.0293	0.1687	0	1
Free other	27	0.0370	0.1925	0	1
Free sth. excluding spray gang	1802	0.1260	0.3319	0	1
Tree age - average	1399	13.8099	6.3538	0	24
Hybrid trees - share	2014	0.4865	0.4588	0	1
Seeds replacing - share	735	0.4080	0.4678	0	1

	count	mean	sd	min	max
Caretaker	2040	0.4049	0.4910	0	1
Paid labour usage	1876	0.8694	0.3370	0	1
Used fertilizer	2039	0.4129	0.4925	0	1
Used insecticides	2039	0.8141	0.3891	0	1
Used fungicides	2039	0.3551	0.4787	0	1
Used spray m. non-gov.	2039	0.6974	0.4595	0	1
Used gov. spray gang	2035	0.8226	0.3821	0	1
Used other inputs	2138	0.0557	0.2293	0	1
Event black pod	2138	0.1969	0.4162	0	3
Event swollen shoot	2138	0.0491	0.2162	0	1
Event caterpillar	2138	0.0496	0.2193	0	2
Event other insects	2138	0.0486	0.2152	0	1
Event other	2138	0.1235	0.3497	0	2
Event number	2138	0.5744	0.7772	0	4

Table A.4.: Mean levels of the factors of production and output depending on credit and saving and their standard deviations

	Cocoa	Land	Labour	Inputs	Capital
borrowed	1708.9	17.57	385.0	301.7	199.63
	(2643.1)	(18.51)	(474.1)	(502.0)	(378.40)
did not borrow	1764.4	17.59	330.8	327.8	216.20
	(2389.2)	(19.37)	(497.2)	(930.0)	(393.27)
saved	2401.3	21.13	457.8	446.7	273.96
	(3183.8)	(23.12)	(641.9)	(1115.4)	(470.40)
did not save	1389.3	15.56	278.0	257.0	179.32
	(1829.6)	(16.27)	(370.6)	(700.0)	(334.53)