Technical efficiency, technical change and return to scale of rice, maize and agricultural production in Vietnam

Dissertation

to obtain the Ph. D. degree

In the International Ph.D. Program for Agricultural Sciences in Göttingen

(IPAG)

in the Faculty of Agricultural Sciences,

Georg-August-University Göttingen, Germany

presented by

Tran, Duc Tri

born in Ha Noi, Vietnam

Göttingen, March 2019

D7

1. Name of supervisor: Prof. Dr. Xiaohua Yu

2. Name of co- supervisor: Prof. Dr. Bernhard Brümmer

3. Name of further member of the examination committee: Prof. Dr. Jan-Hanning Feil

Date of dissertation: …………………….

Summary

The rapid growth of human population has impacted on global food security, development and health. Therefore, efforts to achieve the sustainable development goals of reducing poverty and hunger needs to particularly focus on the critical linkages between agriculture, nutrition, health and poverty reduction. Agriculture plays a significant role in largely agro-dependent developing economies as a source of livelihood to the rural population, foreign exchange earner and source of food to the growing populations. Recent statistics estimate the world population at 7.6 billion by mid 2017, an estimate projected to double by year 2050. Interestingly, about 60 percent of the population is from Asia and largely reside in rural areas with farming as main occupation. This put pressure on land and other natural resources to feed the growing population amid dwindling rich arable lands as a result of rapid urbanization. The agriculture sector in many of the Asian countries, just as in other developing countries is challenged by increased land fragmentation and dwindling productivity trends over the years.

Vietnam, one of the Southeast Asian countries, continues to face these difficulties in the agricultural sector. The country however has emerged from the challenges of food security as a net importer of major staples in the 1980s such as rice and maize to net exporters of various agricultural commodities, courtesy of the ‘Doi Moi’ revolution of 1986 that changed Vietnam from a centrally planned to open market economy. This was accompanied by institutional changes facilitated by the Directive 100CT/TW of 1981 and the Resolution 10 of 1988 that transformed Vietnam’s agriculture and related sectors. The changes ushered in policies that initiated significant structural transformation in the sector that saw labour movement out of agriculture to the feeder industries. By 2016, the proportion of labour in agriculture had fallen from 63 percent at the turn of the millennium to 42 percent.

Despite the structural transformation, agriculture still contributes significantly to Vietnam’s GDP at 15 percent, as a major source of employment to largely rural population and as a source of food security. Two major agricultural commodities stand out in Vietnam agriculture, rice and maize as staples and major sources of food security, incomes among farm households and raw materials to processing industries. The crops grow on relatively small yet increasingly fragmented pieces of

lands of about 1.5 ha on average. This has had major implications on agricultural production in terms of efforts to increase yields, productivity and efficiency in the small farm sector that contributes more than 70 percent of aggregate agricultural production.

Both crop and livestock productions are also significantly affected by climate change and land scarcity in Vietnam. Despite significant investments to improve agricultural production and productivity in Vietnam over the years, crop and livestock productivity and efficiency in production, as well as returns to scale on agriculture remain low. While many studies have looked at these aspects in agricultural production in many developing countries, analysis of technical efficiency (TE) and overall returns to scale on agriculture has received much less attention particularly in terms of innovative approaches to analyze efficiency using panel data. This is even more pronounced in the context of Vietnam where return to scale, TE and drivers of technical inefficiency are scarcely studied using panel data.

The dissertation seeks to fill the knowledge gaps using the case of Vietnam by analyzing the returns to scale in Vietnam agriculture (crop and livestock production), TE in rice and maize production and drivers of technical inefficiency in the production of agriculture sector and particularly these two crops. We innovatively employ a combination of stochastic frontier distance functions, stochastic frontier and Tobit models on 5-waves panel data from smallholder farmers between 2008 and 2016. The data comprised Vietnam Access to Resources Household Surveys (VARHS).

The dissertation is an amalgamation of three related papers on the aforementioned topics as follows. Paper 1 presents findings from analysis of TE, technical change, and return to scale in Vietnam Agriculture in the period 2008-2016 as well as identifying the factors affecting the technical inefficiency using four models of output distance functions and Tobit on panel data of 487 households in each of the five rounds of survey. The findings show that the level of TE of Vietnam agricultural is 89.29%, of which, the highest belongs to Lao Cai province, followed by Lai Chau and the lowest is Phu Tho province. The average technical change for the whole study period tends to decrease by 4.43%. The result of elasticity estimation indicates that all inputs take positive impact on increasing the value of agricultural output in Vietnam. In which, land plays the biggest role, followed by intermediate cost and labour. Return to scale is estimated to be 78.49% and tends to increase during the study period. The study also shows that the ethnicity of household head effects positively on TE while the number of household members and land fragmentation negatively influence TE

The second paper focuses on TE in rice production using stochastic frontier models on panel data of 1555 households in each of the five rounds of survey and investigate the drivers of TE using Tobit model. Log-likelihood ratio (LR) test is used to select the optimal model. The results show that TE score of Vietnam is 92.62% and increases over the study period. There are TE differences among the six economic sectors, the highest is the North Central Coast, followed by the Red River Delta and the lowest is the Central Highland. The results of the output elasticity estimate indicate that all inputs positively influence the value of rice production, with hectarage under rice being is the most significant, followed by other inputs such as fertilizer, seed, labour, other costs and pesticide, herbicide. Return to scale is 92% and tends to increase during the study period. Analysis of technical change shows that the production frontier function of rice production tends to increase 1.06% in each period. From 2006 to 2016, TFP growth is 4.29%. The results also show that the gender and level of education (most educated) of the household head, irrigation and land fragmentation index positively and significantly influence TE whereas ethnicity of household head negatively influences TE.

The third paper analyzes TE and risk in maize production in the North Eastern of Vietnam and related drivers using a combination of Just and Pope's stochastic production and stochastic frontier models. A balanced panel data collected every 2 years from 2008 to 2016 among 435 maize households is used. Similarly, LR test is used to select the most optimal model. Results from marginal output risk analysis indicate that land, labour, pesticide and herbicides increase the likelihood of output variances, while seed, fertilizer and other costs reduce the variances. Gender of household head and household size positively and significantly influence TE in maize production. Contrary to the findings on rice, gender and level of education (most educated) of the household head, irrigation and disaster indices negatively influence TE. The average TE of maize production in the North West is 82.75% and increased steadily over the period and by 1-2%. The highest TE is observed in Dien Bien province, followed by Lao Cai and the lastly is Lai Chau. Research gives some recommendations to increase maize production, to eliminate technical inefficiencies, and to minimize the impact of risk during production.

Acknowledgements

First of all, I would like to express my sincere gratitude to my first supervisor, Prof. Dr. Xiaohua Yu for his patience, motivation, and the knowledge imparted during the PhD study. I also thank the rest of my thesis committee members, Prof. Dr. Bernhard Brümmer and Prof. Dr. Jan-Hanning Feil for their invaluable comments during the doctoral seminars that improved my work. I acknowledge the immense support from all the GlobalFood Research Training Group (RTG1666) colleagues and the Chair of Agricultural Economics in Developing and Transition at the University of Göttingen.

I am grateful to the Vietnamese Ministry of Education and Training (Steering Committee of Program 911) for the financial support to successfully undertake the doctoral program in Germany. I would like to thank the Ministry of Agriculture and Rural Development, the Vietnam National University of Agriculture, the Faculty of Economics and Rural Development, the Department of Economics for allowing me to enroll and complete my studies. Special thanks Prof. Dr. Duc Vien Tran, Assoc. Prof. Dr. Dinh Thao Tran, Dr. Tat Thang Nguyen for their unwavering support throughout the period.

I acknowledge the support from the Institute of Labour Science and Social Affairs under the Ministry of Labour Invalid and Social Affairs of Vietnam, that provided the Vietnam Access to Resources Survey data sets used in this work, I would like to thank the leaders of the Institute of Labour Science and Social Affairs, especially Dr. Nguyen Thi Lan Huong and M.A Le Ngu Binh who facilitated the process. Special thanks to my colleagues, MA. Ha Thi Thu Huong and doctoral candidate, Nguyen Tran Minh Thanh for the stimulating discussions and sleepless nights we worked together in data processing.

To my Vietnamese and international friends, and members of the Vietnamese Students Association in Göttingen, thank you very much for making my stay in Göttingen more meaningful, lively and memorable. Your company will forever be most appreciated.

Last but not the least, this work would have not been possible without the unrelenting spiritual support from my parents and sister in my life generally and more so during my studies. To my wife and daughter, thank you for the company and support in all circumstances, sharing your love and sympathy throughout this journey.

List of Abbreviations

CD	Cobb-Douglas
CH	Central Highland
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
GSO	General Statistics Office (Vietnam)
ha	Hectare
kg	Kilogram
m^2	Metres squared
MRD	Mekong River Delta
NCC	North Central Coast
NE	North East
NW	North West
RRD	Red River Delta
RTS	Return to Scale
SC	Scale effect
SCC	South Central Coast
SFA	Stochastic Frontier Analysis
TC	Technical change
TE	Technical efficiency
TFE	True Fixed Effect
TFP	Total factor productivity
TI	Technical inefficiency
TRE	True Random Effect
VARHS	Vietnam Access to Resources Household Survey
VHLSS	Vietnam Household Lliving Standards Survey
VND	Vietnamese Dong
VRS	Variable Returns to Scale

List of contents

List of tables

List of figures

1. General introduction

1.1. Background

The world population reached 7.6 billion in mid-2017 and is projected to reach 9.6 billion by year 2030 (PRB, 2017). About 60 percent of the 7.6 billion are living in Asia. The rapid population growth led to increased urbanization and raised pressure on land for agricultural production and the dwindling natural resources in the face of climate change (Kirchmann & Thorvaldsson, 2000). In the context of rising food demand and declining agricultural production, the improvement of agricultural productivity and efficiency is the key to sustaining food production and food security (Devendra, 1999). Therefore, the increase in technical efficiency, change and return to scale in agriculture in general and in the production of major staple foods in particular has received great attention.

Vietnam is one of the Asian highly populated countries with population estimated at 95.54 million and is facing challenges of agricultural production, especially for staple foods such as rice and maize. This has been exacerbated by climate change, declining productivity trends, increased land fragmentation and reduced soil quality. After independence and unification in 1975, the country has increasingly made institutional reforms to foster economic growth and development as well as ensure food security. In particular, reforms of 1986 "Doi Moi" resulted in a shift from centrally planned to a multi-component commodity economy, operating under the market mechanism with the State's management under the socialist orientation. The reforms have led to significant agricultural investments in productivity increasing technologies to boost agriculture and changed Vietnam from a net importer to net exporter of major agricultural commodities such as rice, coffee, pepper and cashew nuts both globally and in the region.

Given that about 60% of Vietnam's population reside in rural areas and mainly depend on agriculture as a source of livelihood, such investments have simultaneously contributed to poverty reduction, increased employment, incomes and food security, thereby ensuring social stability and economic development. (De Janvry & Sadoulet, 2009; Irz, Lin, Thirtle, & Wiggins, 2001; Kassie, Shiferaw, & Muricho, 2011; Thirtle, Lin, & Piesse, 2003). The increased structural transformation has also led to the movement of labour out of agriculture to feeder industries, trade and services.

Normally, measuring TE begins with a description of production technology and may be different but equally the same approach given that their results from different approaches will converge (Kalirajan & Shand, 1999). SFA as an approach, originated from Meeusen and Van Den Broeck (1977); Aigner, Lovell and Schmidt (1977); and Battese and Corra (1977). Battese and Coelli (1988) defined the TE of a given firm at a given time period as the ratio of its mean production (conditional on its levels of factor inputs and firm effects) to the corresponding mean production if the firm utilized its levels of inputs most efficiently. Improving the specification SFA model may be related to changing productivity with variation of inputs, which is useful for policy makers in developing countries in developing policies to improve productivity and management in general (Shapiro & Müller, 1977).

In agriculture, the heterogeneity of climate and agroecology presents diverse opportunities for various farming systems involving crop and livestock productions. The variety of outputs often creates high competitive value. Therefore, a reasonable combination of inputs and outputs can increase profitability (income) for households, create jobs, reduce poverty and preserve scare resources such as land and water resources (Joshi, Gulati, Birthal, & Tewari, 2004; Lemaire, Franzluebbers, de Faccio Carvalho, & Dedieu, 2014; Pingali & Rosegrant, 1995; Ryan & Spencer, 2001). There is a diversity of climate and topography in the Northernt Vietnam. Furthermore, agricultural production is usually implemented in small scale by households. Thus, it is a favorable condition for diversifying crops and livestock.

In the process of industrialization and modernization of Vietnam, there has been a significant shift from agriculture to industries, trade and services due to the low opportunity cost in long production cycles coupled with production and marketing risks. In 2016, the value of agricultural products accounted for 14.57% of the total value of domestic products. This puts great pressure on the agricultural sector in the face of the need to improve technology, productivity and efficiency to sustain growth. This has also necessitated on-farm diversification mixes of crop and livestock enterprises to diversify production risks. Distance function was applied to measure TE with multiple-outputs based on the concepts of radial contraction and expansion (Malmquist, 1953; Shephard, 2012). It is assumed that not all the firms achieve TE, SFA is therefore useful in estimating parametric stochastic frontier specification of distance function (T. Coelli & Perelman, 1996, 1999, 2000; O'Donnell & Coelli, 2005). There have been several researches that used distance function to measure TE in various sectors such as transport (railways) and banking. However, this has been much less explored in agriculture, in exception of the studies on dairy farming in Europe and cassava farming in Nigeria. We fill this gap by analyzing TE using these approaches in the context of Vietnam.

Despite the remarkable progress in enhancing agricultural production, increasing productivity of major staples such as rice and maize remains a major challenge in Vietnam. Rice is the main staple food for more than half of the world's population. In Vietnam, it is the most important crop and is grown in most parts of Vietnam. There are two largest paddy lands in Vietnam, the Red River Delta and Mekong River Delta. In 2016, rice production area reached 7207.4 thousand hectares, productivity was 4.99 tons/ha and a yield of 35942.7 thousand tons of rice. Rice production contributes 40% of total agricultural output. There are some previous studies that used SFA or DEA or both to estimate technical efficiency of rice production in Vietnam. However, there is a lack of studies that found the most appropriate model to predict technical efficiency. Besides, there is lack of time trend in studied models, not provided sufficient evidence of the impact of regional variations on rice production results nor had the level of TE in rice production been established among regions in Vietnam. It is unclear what policies have affected the technical efficiency in rice production in Vietnam.

Maize also contributes to reducing poverty in rural areas, lessening deforestation and land degradation. Maize is the second most important crop in Vietnam, especially for animal feeds. The Northern mountain areas is the agro-ecological area that provides most of Vietnam's maize production. Maize is grown in the northern highlands with an average of 1.5 ha per household, which plays an important role in the household economy. Nevertheless, maize production is usually facing risks, some inputs may increase or decrease output risks. So far, there is no research on the technical efficiency of maize production in the northern mountainous areas of Vietnam, where the largest production in Vietnam has been made. In particular, there has been no serious examination of the risks of production related to the utilization of inputs in maize production in particular and in the agricultural sector in general. Evaluation of the effectiveness of maize farming households in combination with production risk will make TE measurement more accurate, by evaluating the effect of input use on output variance (production risk), it will elicit important policy implications in agricultural development planning (Jaenicke, Frechette, & Larson, 2003; Villano & Fleming, 2006).

This dissertation contributes to the body of knowledge on TE and productivity in the small farm sector in three ways. Firstly, it provides an overall view of Vietnam agriculture in general through applying distance function to study technical efficiency, technical change and return to scale of Vietnam agriculture sector. Concurrently, it proposes suitable recommendations to improve and convert in the conditions and circumstances of Vietnamese households. Secondly, it indicates an overall picture of rice production in the whole country of Vietnam, providing enough evidences of the technical efficiency levels between the study areas as a good base for rice production planning. The technical change and return to scale of Vietnam rice production are also examined. This research proposes recommendations for policy makers to improve technical efficiency of rice production households. Thirdly, the study demonstrates a look at maize production in Vietnam under technical inefficiency and production risk. This study selected model and provided technical efficiency level, technical change and return to scale of maize production in the Northwest of Vietnam. This study shows the impact of inputs to output risk and proposes solutions to minimize technical inefficiency as well as output risk.

1.2. Research objective and dissertation outline

This dissertation becomes urgent to solve the following research questions:

- What is the level of technical efficiency, technical change and return to scale of Vietnam agriculture? Which factors determine the technical efficiency of agricultural production? How to improve income as well as technical efficiency of Vietnam agriculture sector?

- Which model is suitable to estimate technical efficiency of rice farming? How different is the technical efficiency of rice producing households among regions? Which factors determine the technical efficiency of rice production households? How to improve technical efficiency in Vietnam rice production?

- What is the level of technical efficiency, technical change and return to scale of maize producers in Northwestern Vietnam? Which are the drivers of on technical inefficiency? How does input factors affect output risk? How can technical inefficiency be improved and output risks minimized?

The research questions are addressed using the Vietnam Access to Resources Household Survey (VARHS) data managed by the Institute of Labour Science and Social Affairs, under the

Ministry of Labour Invalid and Social Affairs of Vietnam. The VARHS began in 2002 with a sample of 932 households from 4 provinces. In 2006, the sample was expanded to 2324 households in 12 provinces, round 2008 was 3223 households, round 2010 was 3202 households (in which, 2200 panel households), round 2012 was 3700 households, round 2014 was 3648 households and 3582 households in round 2016. All three papers in this dissertation used balanced panel data from 2008-2016.

Apart from the Introduction part in chapter 1- "General Introduction", the remainder of the dissertation is structured as follows:

Chapter 2 presents the first essay entitled "Technical efficiency, technical change, and return to scale in Vietnam Agriculture: A stochastic output distance function approach"

Chapter 3 demonstrates the second essay entitled "Modeling technical efficiency using stochastic frontier production function for panel data: An application in rice farming in Vietnam"

Chapter 4 displays the third essay entitled "Technical inefficiency and production risk of maize farming: A Case study of the Northwestern, Vietnam"

The last chapter concludes by summarizing our main findings, study implications, limitations, and suggestions for further research.

2. Technical efficiency, technical change, and return to scale in Vietnam Agriculture:

A stochastic output distance function approach

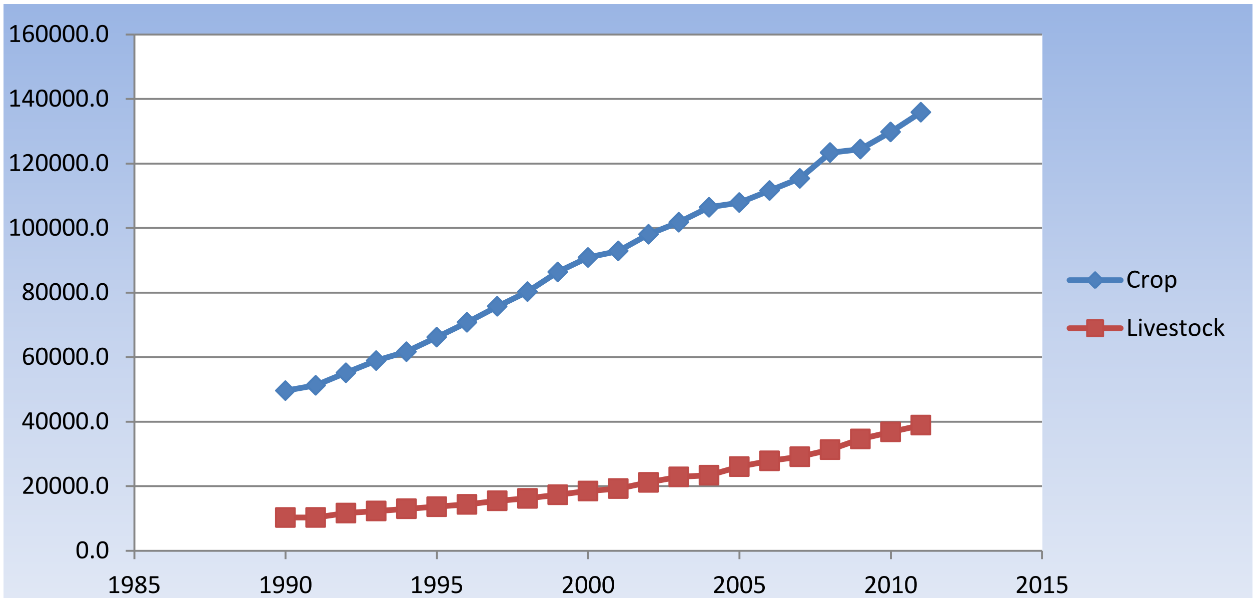
Abstract

A few studies have analyzed technical efficiency and return to scale in Vietnam agriculture using panel data. In this paper, we employ stochastic output distance function approach on panel data collected from 487 households with 5 surveys, to analyze technical efficiency and drivers of technical efficiency, technical change, and return to scale in Vietnam Agriculture. Overall, the results show an 89.29 percent level of technical efficiency, with varying levels of efficiency across provinces, the highest and lowest efficiencies being in Lao Cai (90.51%) and Phu Tho (87.08%) provinces respectively. Interestingly, efficiency is positively and significantly influenced by number of household member and number of plot while ethnicity of household head negatively influenced efficiency. The average technical change decreased by 4.43 percent. The return to scale is 0.78 and tends to increase over time. Average TFP growth during the period reached 2.71%, however, there was a downward trend over the years. Improving technical efficiency, change and return to scale will involve on-farm and off-farm investments by farmers as well as Government to increase agricultural production.

2.1. Introduction

The economic growth in Vietnam in general and of the agricultural sector in particular has been remarkable since the "Doi Moi" institutional reforms of 1986. This is attributed to the two pieces of legislation, *The Directive 100 CT/TW* of 1981 and *The Resolution 10* of 1988 that transformed Vietnam from a food importer to exporter, especially of rice. Vietnam is currently one of the leading globally rice exporters. In addition, it positively contributes to the process of poverty reduction, ensuring food security for the whole country in general and northern Vietnam particular.

Recent statistics show that the total value of production (TVP) in the agricultural sector has significantly increased over the years (Figure 2.1), with crops sub-sector accounting for a larger share of the TVP than livestock subsector. As of 2011, the TVP was 175 billion VND, of which the crops sub sector contributed about 77% (GSO, 2011).



(Source: GSO database www.gso.gov.vn)

Figure 2.1. Total output value of crop and livestock (fixed price of 1994, unit: billion VND)

The institutional reforms further facilitated structural transformation across sectors of the economy in the wake of trade liberalization (open market) policies that further increased industrialization. Consequently, the factors of production (lab, labor, and capital) freely moved across the various productive sectors thereby easing labour from farm to off-farm sectors (industry). The proportion of labour fell from 63% in year 2000 to 42% by 2016 (GSO, 2000, 2016). Since most of the industries were in urban areas, rural to urban migration increased thus putting pressure on arable land as more land was allocated to urban housing. The urban residential land increased sharply from 75,128 hectares (ha) in 2000 to 156,500 ha by 2016, Coupled with rapid population growth, the agricultural sector had to transform to ensure food security and sustain production. This called for significant investment in modern technologies for crop and livestock production to improve efficiency and productivity. Many studies have explored the impacts of the various technologies on employed under various farming systems in Vietnam, owing to the heterogeneous of agro-ecology. This is mainly in terms of technological efficiency of in the various production regimes (Dao & Lewis, 2013; Hoang Linh, 2012; Huynh-Truong, 2009; Khai, Yabe, Yokogawa, & Sato, 2008; Pedroso et al., 2018; Tuan M.Cao, Sutonya Thongra, & Kiatpathomchai, 2017) and animal husbandry (Akter, 2003; Den, Ancev, & Harris, 2007; Jabbar & Akter, 2006; Tung, 2010). Most of these studies used traditional methods of measuring technical efficiency (TE) such as Data Envelopment Analysis (DEA) (non-parametric approach) and Stochastic Frontier Analysis (SFA) (parametric approach). One notable study is Ho's (2012) that analyzed TE and technical change in the agricultural sector in Vietnam between 1990 and 2006 using both DEA and SFA. The study estimated the average TE of 75.3% and 79.3% using the DEA and SFA approaches, respectively. The technical efficiency change was reduced over the years (-1.2 % by DEA method and -3.1% by SFA method). Both methods estimated a technical change of 1.5% annually.

The two traditional methods of measuring TE had two problems. They could not describe scenarios of multi-output technology without price information and the methods could not account for the objective behavior such as profit maximization or cost minimization. A distance function approach overcomes these two drawbacks when measuring the TE and productivity from which the technology can be described. The concept is based on radial contraction and expansion (Malmquist, 1953, Shephard, 2012). However, the distance function is similarly expressed as the DEA and SFA methods and is estimable using both econometric and mathematical programming methods, The underlying assumption is that not all the firms achieve TE. Data envelopment analysis (DEA) is one of the most important methods to measure nonparametric. The DEA involves linear programming hence there is no need to use production technology. On the contrary, SFA requires production technology in the specification of distance function (T. Coelli & Perelman, 1996, 1999, 2000; Hetemäki, 1996; O'Donnell & Coelli, 2005).

Many studies have used distance functions to measure technical efficiency and change and returns to scale. For instance, in measuring TE of European railways (T. Coelli & Perelman, 1999, 2000); in measuring TE, technical change and return to scale in banks (Abdul-Majid, Saal, & Battisti, 2008; Cuesta & Orea, 2002; Feng & Serletis, 2010); in measuring productivity growth of European farms (Brümmer, Glauben, & Thijssen, 2002; Emvalomatis, 2012; Newman & Matthews, 2006, 2007); decomposing energy productivity change in OECD countries (Wang, 2007); decomposing the effects of governance changes on bank efficiency in China (Jiang, Yao, & Zhang, 2009); estimating technical efficiency, input substitution and complementary effects of cassava production in Nigeria (Ogundari & Brümmer, 2011).

It is easy to measure the output factors in non-agricultural sectors such as manufacturing and trading. However, in the agriculture sector, the process is complicated due to the diversity of output factors for crop and livestock enterprises. Brümmer (2006) used a stochastic output distance function with panel data to analyze policy reform and productivity change in Chinese agriculture.

Research indicated that in the first sub-sample, the TE was relatively low and TE changed over time at a small rate. In the second sub-sample, the TE was higher and TE tended to increase.

Despite the extent literature on TE, technical change and return to scale in both agricultural and non-agricultural sectors in many developing countries, empirical studies on this in the context of Vietnam remain limited especially using distance functions. We fill this gap by employing distance functions on panel farm household data to measure TE and drivers of TE, technical change and return to scale in Vietnam agriculture. Tobit model is used to analyze the drivers of TE. The findings have implications on pathways through which technological change could be enhanced to ensure sustainable growth of the agriculture sector in Vietnam as well as other developing countries whose agricultural sector growth is key to livelihood enhancement.

The rest of the paper is organized as follows. The next section presents the theoretical framework, Section 3 describes the methodology of the study whereby data and sources of data and model specifications are outlined. Section 4 discusses the study findings and section 5 concludes.

2.2. Theoretical framework

Distance function is increasingly being used for experimental research. In the case of multiple outputs, the distance function is substituted for the output function with multiple outputs. One of the advantages of the distance function is that it manifests itself simultaneously with the multiple outputs and manufacturing technology of many inputs, so it is easy to analyze TE. When price information is not available or assumptions such as cost minimization or profit maximization are not appropriate, then traditional methods cannot be applied to multi-output production technologies by estimating the cost function and the profit function. Therefore, the distance function is appropriate in such situations.

A production technology can be specified as follows:

$$P^t(X^t) = \{Y^t\text{: Y producible from X}\} \quad (1)$$

where $X \in R_+^K$ and $Y \in R_+^M$ represent input and output vectors at time t=1, 2, ...T

Distance function was first introduced by Shephard (1970) and was defined by the set of output vectors feasible for each input vector X, P (X) as follows:

$$D^{Ot}(Y^t, X^t) = \min.\left\{\Theta > 0 : \frac{Y^t}{\Theta} \in P^t(X^t)\right\} \quad (2)$$

The output distance function is; non-decreasing, positively linearly homogeneous and convex in outputs, and decreasing in inputs (Lovell, Travers, Richardson, & Wood, 1994). There are 2 main characteristics of this function as follows: (i) homogeneity: $D^{Ot}(\omega Y^t, X^t) = \omega D^{Ot}(Y^t, X^t)$ for any ω>0; (ii) $D^{Ot}(Y^t, X^t)$ is less than or equal to 1 if the output vector (Y^t) is an element of the feasible production set of $P^t(X^t)$. That is $D^{Ot}(Y^t, X^t) \leq 1$ if $Y^t \in P^t(X^t)$. The distance function is equal to 1 (unity) if Y is on the boundary of the productive set, it means that $D^{Ot}(Y^t, X^t) = 1$ if $Y^t \in$ Isoquant $P^t(X^t)$.

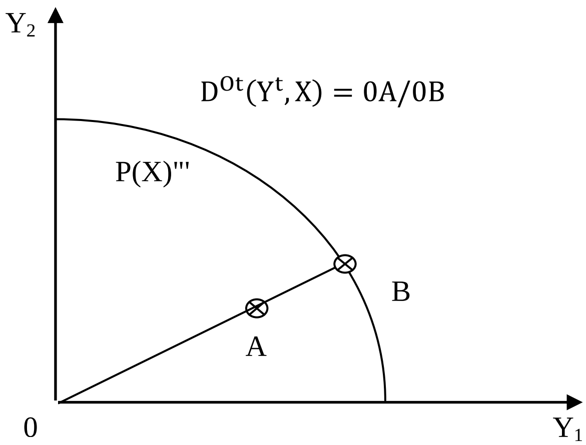


Figure 2.2. Output distance function measure of technical efficiency (M=2)

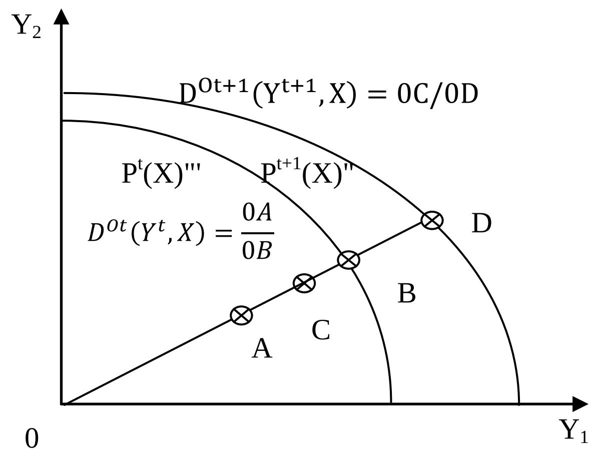


Figure 2.3. Technical change and technical efficiency change (M=2)

In Figure 2.2, the output set P(X) is surrounded by production capacity curve, which describes the TE of production with each output combination given by the input. The output-oriented measure of TE, $TE^O(Y, X)$) overlaps with the output distance function $D^O(Y, X)$.

In Figure 2.3, it is assumed that there are two outputs and inputs are constant. Technical change leads to a change in the set of outputs from $P^t(X)$ to $P^{t+1}(X)$. TE change is measured by the ability of the manufacturer to improve the efficiency from t to period t + 1, which is the change from $D^{Ot}(Y^t, X)$ to $D^{Ot+1}(Y^{t+1}, X)$.

Distance functions can be estimated using various methods namely: corrected ordinary least squares (COLS) regression, DEA, SFA and parametric deterministic linear programming (PLP). The output distance function can be written as follows, $\ln D_{it}^{O}(Y, X, t) \leq 1$. In the context of TE, it takes the form: $\ln D_{it}^{O}(Y, X, t)\exp(U) = 1$, where $U \geq 0$ is output-oriented TE.

By totally differentiating of both sides:

$$\sum_m \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth \ln Y_m}\dot{Y}_m + \sum_k \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth \ln X_k}\dot{X}_k + \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth t} + \frac{\eth U}{\eth t} = 0 \qquad (3)$$

$$\sum_m \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth \ln Y_m}\dot{Y}_m + \sum_k \frac{(\eth \ln D_{it}^{O}(Y,X,t)/\eth \ln X_k)\ \ \sum(\eth \ln D_{it}^{O}(Y,X,t)/\eth \ln X_k)}{\sum(\eth \ln D_{it}^{O}(Y,X,t)/\eth \ln X_k)}\dot{X}_k$$

$$+ \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth t} + \frac{\eth U}{\eth t} = 0 \qquad (3a)$$

we define $S_k = \frac{W_k X_k}{\sum_n W_k X_k} = \frac{\eth \ln D_{it}^{O}/\eth \ln X_k}{\sum_n \eth \ln D_{it}^{O}/\eth \ln X_k}$; $RTS = -\sum_k \frac{\eth \ln D_{it}^{O}(Y,X,t)}{\eth \ln X_k}$;

we can rewrite equation (3a) as follow

$$=> \qquad \sum_m R_m \dot{Y}_m - \sum_n S_k \dot{X}_k\,(RTS) + {}^{\eth \ln D_{it}^{O}}/_{\eth t} + {}^{\eth U}/_{\eth t} = 0 \qquad (3b)$$

$$=> \qquad \sum_m R_m \dot{Y}_m = \sum_n S_k \dot{X}_k\,(RTS) - {}^{\eth \ln D_{it}^{O}}/_{\eth t} - {}^{\eth U}/_{\eth t} = 0 \qquad 4$$

We define:

$$=> \qquad T\dot{F}P = \sum_{m=1}^{M} R_m \dot{Y}_m - \sum_{k=1}^{N} S_k \dot{X}_k \qquad (5)$$

substituting (4) to (5)

$$=> \qquad T\dot{F}P = (RTS - 1)\sum_{n=1}^{N} S_k \dot{X}_k - {}^{\eth \ln D_{it}^{O}}/_{\eth t} - {}^{\eth U}/_{\eth t} \qquad (5a)$$

where: a dot over variable indicates the respective growth rate

$$R_m = \frac{P_m Y_m}{\sum_n P_n Y_n} = \frac{\eth lnD_{it}^{O}/\eth lnY_m}{\sum_n \eth lnD_{it}^{O}/\eth lnY_m} = \frac{\eth lnD_{it}^{O}}{\eth lnY_m}$$

Because of the homogeneity property of $D_{it}^{O}, \sum_n \eth \ln D_{it}^{O} / \eth \ln Y_m = 1$, uniformity is then used to estimate the model specified in equation 5, and the output distance functions are written as follows

$$D_{it}^{O}(Y_{mit}, X_{kit}, t)/Y_{Mit} = D_{it}^{O}(Y_{mit}/Y_{Mit}, X_{kit}, t) \quad (6)$$

$$<=> \quad \ln D_{it}^{O}(Y_{mit}, X_{kit}, t) - \ln Y_{Mit} = \ln D_{it}^{O}(Y_{mit}/Y_{Mit}, X_{kit}, t) \quad (7)$$

$$=> \quad -\ln Y_{Mit} = \ln D_{it}^{O}\left(\frac{Y_{mit}}{Y_{Mit}}, X_{kit}, t\right) + U_{it} \quad (\text{where } \ln D_{it}^{O}(Y_{mit}, X_{kit}, t) = -U_{it} \leq 0) \quad (8)$$

The second component of equation 5 shows technical change (technical progress), which can be positive or negative.

$$\eth \ln D_{it}^{O}(Y_{mit}, X_{kit}, t)/\eth t = \ln D_{it}^{O}\left(\frac{Y_{mit}}{Y_{Mit}}, X_{kit}, t\right)/\eth t \quad (9)$$

The last component of equation 5 shows the effect of technical efficiency change: $\eth U/\eth t$. In which, U_{it} is estimated by model $-\ln Y_{Mit} = \ln D_{it}^{O}\left(\frac{Y_{mit}}{Y_{Mit}}, X_{kit}, t\right) + U_{it}$. Thus, through the estimation of the output distance function, all three components of TFP growth can be obtained.

2.3. Methodologies

This section first describes the data used in analyzing TE, technical change and returns to scale in Vietnam agriculture before deriving the equations to measure these indicators and finally the model specifications used in generating the results.

2.3.1. Data and sources

Currently, there are two main nationally representative survey data sets in Vietnam namely, the Vietnam Household Living Standards Survey (VHLSS) collected by the by the General Statistics Office with support from the World Bank since 1992 and the Vietnam Access to Resources Household Survey (VARHS), implemented by the Institute of Labour Science and Social Affairs (Ministry of Labour, Invalids and Social Affairs, Viet Nam), funded by Danish International Development Agency (DANIDA). The VARHS has observations from farm households over time while VHLSS is a census that targets all urban and rural households.

This study uses the VARHS data. The VARHS survey began in 2002 with a small sample (932 households in 4 provinces). In 2006, the sample was increased to 2324 households in 12 provinces, round 2008 was 3223 households, round 2010 was 3202 households (in which, 2200 panel households), round 2012 was 3700 households, round 2014 was 3648 households and 3582 households in round 2016. We used a balanced panel of 487 farm households interviewed during the five survey rounds between 2008 and 2016, giving a total of 2435 observations from the three provinces of Vietnam (Table 2.1). Figure 4 presents a map of the study areas. Vietnam Consumer Price Index (CPI) was used to deflate prices using 2008 as the base year to standardize monetary values of inputs and outputs in the production process.

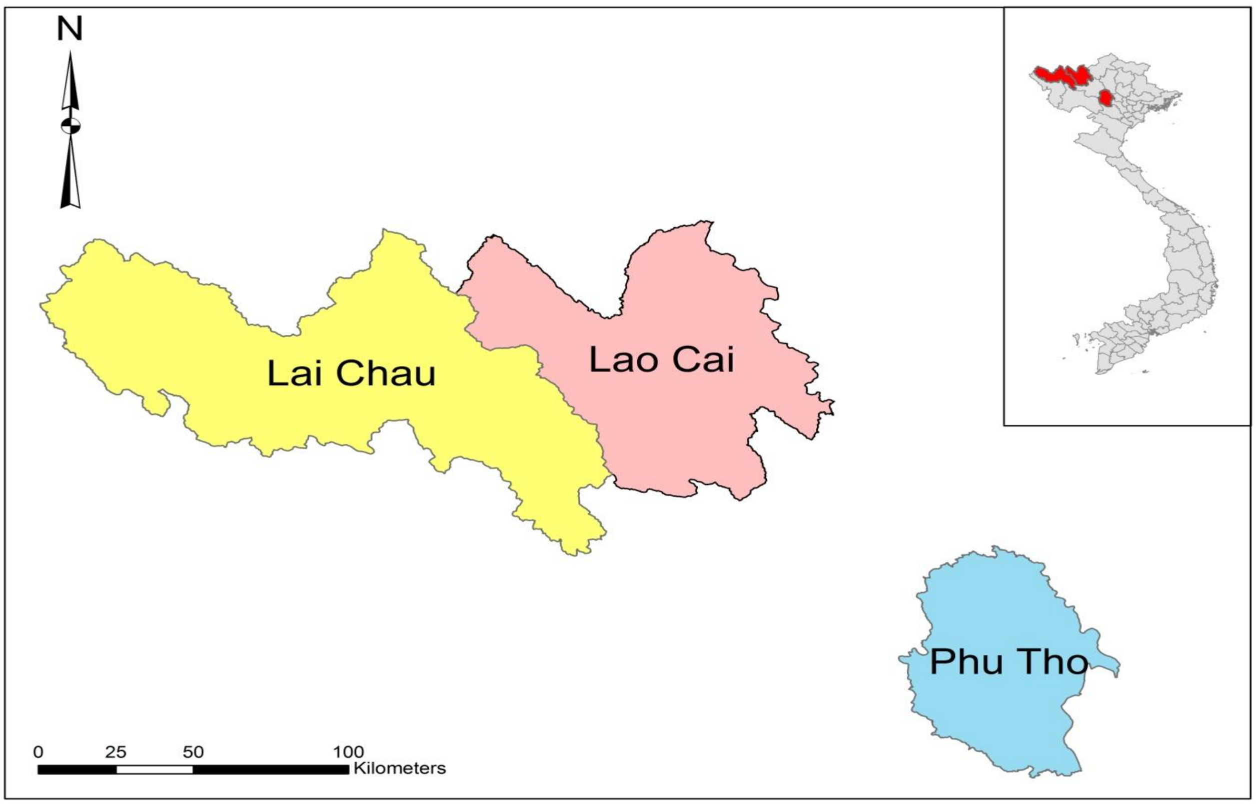


Figure 2.4. Map of research areas

Table 2.1. Observations by regions

Detail	Phu Tho province	Lao Cai province	Lai Chau province
Number of households	112	230	145
Observation	560	1150	725

The descriptive statistics reveal that households with 2 output elements in the agricultural sector (crop and livestock) are concentrated mainly in northern Vietnam. This is plausible for the following reasons: (i). agro-ecological diversity in the north; (ii). large fragmentation of land thus, it is easy to produce a variety of agricultural products; (iii). greater farm diversification that limits scale of crop and livestock production, marketable surpluses and farm incomes.

2.3.2. Specifying the translog output distance function

Assuming a typical production process involving M outputs and K inputs, the translog output distance function can be written as follows:

$$
\begin{aligned}
\ln D_{it}^{O} = {} & \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln Y_{mit} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln Y_{mit} \ln Y_{nit} \\
& + \sum_{k=1}^{K} \beta_k \ln X_{kit} + \frac{1}{2}\sum_{k=1}^{K}\sum_{l}^{K} \beta_{kl} \ln X_{kit} \ln X_{lit} + \sum_{m=1}^{M}\sum_{k=1}^{K} \gamma_{mk} \ln Y_{mit} \ln X_{kit} \\
& + \delta_t t + \frac{1}{2}\delta_{tt} t^2 + \sum_{m=1}^{M} \delta_{tym} t \ln Y_{mit} + \sum_{k=1}^{K} \delta_{txk} t \ln X_{kit} \qquad (10)
\end{aligned}
$$

where "^{O}" explains an output-oriented distance function; D^{O} is not able to observe the value of distance function; $\ln Y_{mit}$ is output (for m=1,2..M) of the firm i in the sample (for i=1,2...N) by the time t (for t=1,2,...T); $\ln X_{kit}$ is input (for k=1,2...K) of the firm i in the sample (for i=1,2...N) by the time t (for t=1,2,...T)

An output function must satisfy certain limits (Coelli & Perelman, 2000; Feng & Serletis, 2010; Yao & Jiang, 2007). The restrictions required for homogeneity of degree one in outputs are:

$$
\sum_{m=1}^{M} \alpha_m = 1; \text{ and } \sum_{m=1}^{M} \alpha_{mn} = 0 \ \ (m = 1,2, \ldots M); \ \sum_{k=1}^{K} \gamma_{mk} = 0 \ \ (k = 1,2, \ldots K)
$$

and restriction required on symmetry are

$$\alpha_{mn} = \alpha_{nm} \ (m, n = 1,2, \dots M); \text{and } \beta_{kl} = \beta_{lk} \ (k, l = 1,2, \dots K)$$

The distance function above cannot be directly estimated, so the distance function by one of the outputs can be made by imposing homogeneous constraints (Lovell et al., 1994; O'Donnell & Coelli, 2005). The characteristic of homogeneity is explained that $D^O(\omega Y, X, t) = \omega D^O(Y, X, t)$ for any ω>0, it can be satisfied by normalizing the output through the use of any output. The M^{th} output is selected for normalization and assigns $\omega = 1/Y_M$. It leads to the following expression $lnD^0\left(\frac{Y_{mit}}{Y_{Mit}}, X_{it}, t\right) = ln[D^0(Y_{mit}, X_{it}, t)/Y_{Mit},]$. After re-arranging and replacing $\ln D^O$ with u, in addition, plus error term to account for random noise. The right-hand side is like a standard stochastic production frontier model.

For convenience, it can be rewritten as follows:

$$\ln\left(\frac{D_{it}^o}{Y_{Mit}}\right) = TL\left(\frac{Y_{mit}}{Y_{Mit}}, X_{it}, t, \alpha, \beta, \gamma, \delta\right) \quad (i = 1,2, \dots N; m = 1,2, \dots M; t = 1,2, \dots T) \tag{11}$$

or

$$\ln(D_{it}^o) - \ln(Y_{Mit}) = TL\left(\frac{Y_{mit}}{Y_{Mit}}, X_{it}, t, \alpha, \beta, \gamma, \delta\right) \quad (i = 1,2, \dots N; m = 1,2, \dots M; t = 1,2, \dots T) \tag{12}$$

Inferably:

$$-\ln(Y_{Mit}) = TL\left(\frac{Y_{mit}}{Y_{Mit}}, X_{it}, t, \alpha, \beta, \gamma, \delta\right) - \ln(D_{it}^o) \quad (i = 1,2, \dots N; m = 1,2, \dots M; t = 1,2, \dots T) \tag{13}$$

Details:

$$-\ln Y_{Mit} = \alpha_0 + \sum_{m=1}^{M-1} \alpha_m \ln\frac{Y_{mit}}{Y_{Mit}} + \frac{1}{2}\sum_{m=1}^{M-1}\sum_{n=1}^{M-1} \alpha_{mn} \ln\frac{Y_{mit}}{Y_{Mit}} \ln\frac{Y_{nit}}{Y_{Mit}}$$
$$+ \sum_{k=1}^{K} \beta_k \ln X_{kit} + \frac{1}{2}\sum_{k=1}^{K}\sum_{l}^{K} \beta_{kl} \ln X_{kit} \ln X_{lit} + \sum_{m=1}^{M-1}\sum_{k=1}^{K} \gamma_{mk} \ln\frac{Y_{mit}}{Y_{Mit}} \ln X_{kit}$$

$$+\delta_t t + \frac{1}{2}\delta_{tt}t^2 + \sum_{m=1}^{M-1}\delta_{tym} t \ln\frac{Y_{mit}}{Y_{Mit}} + \sum_{k=1}^{K}\delta_{txk} t \ln X_{kit} + U_{it} + V_{it} \qquad (14)$$

where V_{it} is random errors and is assumed as independent and identically distributed N(0, σ^2_v), U_{it} is assumed as independent and identically distributed non-negative truncations of the N(μ,σ^2) distribution or half-normal distribution N(0,σ^2). In which, $\frac{Y_{mit}}{Y_{Mit}}$ is the output distance function converted by one of the outputs and Y_{Mit} in the model has been imposed by linear homogeneity properties. Through the characteristics of linear homogeneity properties, the output ratios can be assumed to be exogenous when the output distance functions are defined to radically extend all outputs using a given level of input. Thus, Equation 11 can be transformed into an estimable regression model (Brümmer, Glauben, & Lu, 2006; Brümmer et al., 2002; T. Coelli & Perelman, 2000; Cuesta & Orea, 2002).

For estimation purposes, the negative sign on the dependent variable can be ignored (i.e., use *ln y2* rather than *–ln y2*). TE, technical change and return to scale can be estimated as follows:

$$TE = \exp(-U_{it}) \qquad (15)$$

$$TC = -\frac{\delta \ln D^O_{it}(Y, X, t)}{\delta t} = -(\delta_t + \delta_{tt}t + \sum_{m=1}^{M-1}\delta_{tym}\ln\frac{Y_{mit}}{Y_{Mit}} + \sum_{k=1}^{K}\delta_{txk}\ln X_{kit}) \qquad (16)$$

$$RTS = -\sum_k \frac{\delta \ln D^O_{it}(Y,X,t)}{\delta \ln X_k} \qquad (17)$$

2.3.3. Estimating Technical efficiency using panel data

Using equation 8, multiply both sides by (-1) and add error term V_{it} with the general form of the output distance function as follows:

$$\ln Y_{Mit} = -\ln D_{it}^{O}\left(\frac{Y_{mit}}{Y_{Mit}}, X_{kit}, t\right) - G(t)U_i + V_{it} \qquad (18)$$

where V_{it} is error term and is assumed as independent and identically distributed N(0, σ^2_v) and U_i is time-invariant and assumed as independent and identically distributed non-negative truncations of the N(μ,σ^2) distribution or non-negative and has a half-normal distribution N(0,σ^2).

To study TE changes over the time (U_{it}), there are different assumptions about G(t) (Battese & Coelli, 1992; Cornwell, Schmidt, & Sickles, 1990; Kumbhakar, 1990). This study uses the assumption of Battese and Coelli (1992). They proposed a time-varying model for the efficiency of the period $U_{it} = \{\exp[-\eta(t - T)]\}Ui$ where η is a parameter to be estimated.

2.3.4. Model specification

Translog distance function was estimated by 2 outputs and 3 inputs. The maximum likelihood estimates (MLE) of the parameters of the stochastic frontier production function was obtained by using the Stata frontier function for the different models. In this study, the Battese and Coelli 1992 model was used to run 6 different models based on different assumptions for analysis, In detail: (i) model 1 all parameters are estimated, this is a time-varying decay model and U_i is independent and identically distributed non-negative and has a truncations normal distribution (Ui~ i.i.d N(μ,σ_u^2)); (ii) model 2 assumes that Mu = 0, this is a time-varying decay model, and U_{it} distribution is independent and identically distributed non-negative and has a half-normal distribution (Ui~ i.i.d N(0,σ_u^2)); (iii) model 3 assumes that Eta=0, this is a time-invariant and Uit distribution is independent and identically distributed non-negative truncations of the N (μ, σ_u^2) distribution; (iv) model 4 assumes that Mu = Eta = 0, this is a time-invariant and U_i is independent and identically distributed non-negative and has a half-normal distribution (Ui~ i.i.d N(0,σ_u^2)); (v) model 5 assumes that Mu = Eta = Gamma 0, this is a time-invariant model, and U_i distribution is independent and identically distributed non-negative and has a half-normal distribution (Ui~ i.i.d N(0,σ_u^2)) and U_{it} is absent from the model (the observation variables are full technically efficient).

Using equation 14, the stochastic production frontier with the translog form as follows llivestockneg = α_0+ $\alpha_1$lcropolivestock + α_{11}lcropolivestock_sq + $\beta_1$lland + $\beta_2$llabor + $\beta_3$linter + β_{11}lland_sq + β_{22}llabor_sq + β_{33}linter_sq + β_{12}llandllabor + β_{13}llandlinter + β_{23}llaborlinter + γ_{11}lcropolivestocklland + γ_{12}lcropolivestockllabor + γ_{13}lcropolivestocklinter + $\delta_1$yr + δ_{11}yr2 + $\delta_{ty1/y2}$yrlcropolivestock + δ_{tx1}yrlland + δ_{tx2}yrllabor + δ_{tx3}yrlinter + U_{it} + V_{it}

Where l denotes the natural logarithm of variables, yr denotes a time trend, V_{it} is i.i.d. $N(0, \sigma^2{}_v)$, U_{it} i.i.d. $N(\mu,\sigma^2)$. Output variables include: total turnover of crop production (crop) and total turnover of livestock (livestock). Input variables include: total area used for cultivation and livestock (land), total labour used for cultivation and livestock (labor), intermediate input costs (inter).

Log-likelihood ratio (LR) test was used to find the appropriate model for the data. On the basis of that model, TE, output elasticity, return to scale and technical change are estimated. Tobit model is used to determine the drivers of TE as follows:

TE = ψ_0 + $\psi_1$gender + $\psi_2$family_mem + $\psi_3$ethnic + $\psi_4$land_frag

Where: (i) gender represents the gender of household head (1: male; 2: female); (ii) family_mem shows number of family members; (iii) ethnic represents ethnicity of household head (1: if Kinh or Hoa; 0: otherwise); (iv) land_frag represents number of plot of the household. The next section discusses the study findings, beginning with the descriptive statistics then the model results.

2.4. Results and discussion

2.4.1. Descriptive statistics

Table 2.2 presents the summary statistics of the variables used in estimating the distance function. The total value of crop produced accounts for a high share of the value of agricultural production in northern Vietnam. The intermediate cost of production is relatively high resulting in lower profits in northern Vietnam. Most of the households are headed by men. Few of them are Kinh or Hoa, the rest constitute a few ethnic groups.

Table 2.2. Descriptive statistics of the samples

Variables	Unit	Mean	Sd	Minimum	Maximum
Of the sample					
Total turnover of crop	1000VND	13,307.00	9,563.00	250.00	101,023.00
Total turnover of livestock	1000VND	3,782.00	7,028.00	31.25	176,572.00
Land	m^2	10,940.00	24,962.00	0	836,710.00
Labour	Man days	338.50	212.70	0	1,943.00
Intermediate	1000VND	13,832.00	31,367.00	335.00	1.048e+06
Of the Tobit model					
gender	-	1.08	0.27	1	2
family_mem	number	5.43	1.94	1	16
ethnic	-	0.20	0.40	0	1
land_frag	number	6.10	2.50	1	19

(Source: Vietnam Access to Resources Household Survey)

2.4.2. Model Results

2.4.2.1. Parameter estimates

Four models were fitted and hypotheses tested for their suitability as outlined in Table 2.3. Maximum Likelihood Estimator (MLE) was used to estimate the parameters. Model 1 was taken as the root to compare with other models with different assumptions in order to find the most suitable model:

(i) In comparison with model 2 (appendix 2.1): the LR test rejected the null hypothesis (H_0: $\mu = 0$), therefore the model assuming distribution of U_i was truncated-normal $U_i \sim N(\mu, \boldsymbol{\sigma_u^2})$ is more appropriate than the model assuming distribution of U_i is half-normal ($U_i \sim N(0, \boldsymbol{\sigma_u^2})$), thus model 1 is better than model 2;

(ii) In comparison with model 3 (appendix 2.2): The LR test rejected the null hypothesis (H_0: $\eta = 0$), indicating that the model assuming is time-varying decay was more suitable than model with the assumption is time-invariant, TE of regions can increase or decrease exponentially depending of the sign of the decay parameter η: when $\eta > 0$ it implies that the degree of inefficiency decreases over time and vice versa. In other words, model 1 is better than model 3;

(iii) In comparison with model 4 (appendix 2.3), the LR test rejected the null hypothesis (H_0: $\mu = \eta = 0$), indicating that the model assuming the distribution of U_i is truncated-normal $U_i \sim N(\mu, \sigma_u^2)$) and time-varying decay was more appropriate than model assuming distribution of U_i is half-normal distribution ($U_i \sim N(0, \sigma_u^2)$) and time-invariant, thus model 1 is better than model 4.

(iv) In comparison with model 5 (appendix 2.4): The LR test results shows that the null hypothesis (H_0: $\gamma = \mu = \eta = 0$) is rejected, that is to say the model assuming distribution of U_i is truncat-normal $U_i \sim N(\mu, \sigma_u^2)$), time-varying decay with the existence of the U_i is more applicable than a model with the assumption that distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), time-invariant and without the existence of the U_i the model (the observed variables are fully technically efficient), thus model 1 is better than model 5.

Table 2.3. Hypothesis test for model specification and statistical assumptions

Null Hypothesis	Model	LR test	df	Prob > chi2	Decision
H_0: $\mu = 0$	Model 1 vs model 2	8.66	1	0.0032	Reject H_0
H_0: $\eta = 0$	Model 1 vs model 3	15.22	1	0.0001	Reject H_0
H_0: $\mu = \eta = 0$	Model 1 vs model 4	15.61	2	0.0004	Reject H_0
H_0: $\mu = \eta = \gamma = 0$	Model 1 vs model 5	76047	3	0.0000	Reject H_0
H_0: $\psi_0 = ... = \psi_4 = 0$	Tobit model	56.57	4	0.0000	Reject H_0

In conclusion, after testing the models by different assumptions, model 1 is considered as the best once. As a result, model 1 is chosen for further analysis

Results of using maximum-likelihood estimates to estimate the parameters of the stochastic frontier production function are show in the table 2.4 (model 1). The coefficient of lland is positive and statistically significant at 95% confidence interval and the coefficient of lland is positive and statistically significant at 99% confidence interval, which explains that land and labour used by households in agricultural production in the study area are appropriate and have a positive impact on the value of agricultural production.

The inclusion of the variable yr and yr2 in the production function is to measure the neutral technical change. Similarly, on the interaction terms between time and other stochastic frontier function is intended to measure the error rate of technical change. The coefficient of yr is statistically significance and positive. This explains that neutral technical change occurs over the period and that technical change is increased at a increasing rate.

On the interaction terms between yr and lland is positive and statistically significant at 99% confidence interval. This shows a technical change in agriculture in Vietnam with land deceleration at a small rate by 2.01% over study periods. Analogously, the periodic interaction coefficient between yr and llabor is positive and statistically significant at 99% confidence interval, which explains the technical change in agriculture with labour deceleration at only 3.42% over study periods. The coefficient estimates on the interaction terms between yr and lx_3 is not statistically significant.

Table 2.4. Maximum likelihood Estimates of stochastic frontier model

llivestockneg	Parameter	Coef	Std. Err.	[95% Conf. Interval]	
lcropolivestock	α_1	1.398605***	0.1248168	1.153968	1.643241
lcropolivestock_sq	α_{11}	-0.08156***	0.0139680	-0.1089367	-.0541832
lland	β_1	0.1388471**	0.0705779	0.000517	.2771771
llabor	β_2	0.5657954***	0.1816445	0.2097788	.921812
linter	β_3	-0.1133193	0.1308865	-0.3698521	.1432135
lland_sq	β_{11}	-0.0755007***	0.0030179	-0.0814157	-.0695857
llabor_sq	β_{22}	-0.0805521***	0.0241866	-0.1279571	-.0331472
linter_sq	β_{33}	0.0240966*	0.0145305	-0.0043826	.0525758
llandllabor	β_{12}	0.0539266***	0.0104289	0.0334864	.0743668
llandlinter	β_{13}	-0.0062084	0.0066439	-0.0192302	.0068134
llaborlinter	β_{23}	-0.0610422***	0.0156235	-0.0916638	-.0304206

llivestockneg	Parameter	Coef	Std. Err.	[95% Conf. Interval]	
lcropolivestocklland	γ_{11}	0.0007014	0.0067975	-0.0126216	.0140243
lcropolivestockllabor	γ_{12}	-0.0698745***	0.0173821	-0.1039428	-.0358062
lcropolivestocklinter	γ_{13}	0.0212605*	0.01181	-0.0018866	0.0444077
yr	δ_1	0.4296292***	0.0879664	0.2572183	0.6020402
yr2	δ_{11}	-0.0128896	0.0104313	-0.0333346	.0075554
yrlcropolivestock	$\delta_{ty1/y2}$	0.002718	0.0082685	-0.0134879	0.018924
yrlland	δ_{tx1}	-0.0201418***	0.0059524	-0.0318083	-0.0084754
yrllabor	δ_{tx2}	-0.0342006***	0.0130752	-0.0598275	-0.0085737
yrlinter	δ_{tx3}	0.0022833	0.0079761	-0.0133495	0.0179162
Constant	α_0	-8.977405***	0.8346644	-10.61332	-7.341493
Lnsigma2		3.943982	2.74664	-1.439334	9.327298
Ilgtgamma		6.149972**	2.752817	0.7545492	11.5454
Mu (μ)		-269.2857	741.9236	-1723.429	1184.858
Eta (η)		-0.2782134***	0.0751705	-0.4255449	-0.1308819
$\sigma_S^2 = \sigma_u^2 + \sigma_v^2$		51.62375	141.7919	0.2370856	11240.72
Gamma (γ)=σ_u^2/σ_S^2		0.997871	0.0058483	0.6801691	0.9999903
σ_u^2		51.51384	141.7919	-226.3932	329.4208
σ_v^2		0.1099069	0.0034434	0.103158	0.1166558

(Source: Vietnam Access to Resources Household Survey)

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

While the time variable in stochastic frontier function captures technical change over time (shifting of the production frontier), in an inefficient equation the time variable captures inefficiency change (changes in the distance of the average unit from the rice production frontier). Indicators Eta (η) in the model is statistically significant at 99%. Eta (η) is negative meaning the technical inefficiency increase over time, and units inward from the technical frontier (technological leave) at a rate of 27.82% per period (two years), in other words, TE decreased by the years.

In the model, σ_S^2 is estimated at 51.62, indicating that variance output is caused by technical inefficiencies and random noise. Gamma (γ) is the variance ratio, explaining the total variation in output from the frontier level of output attributed to technical inefficiency. It was estimated at 0.9978, meaning 99.78% of total variation in the value of agricultural production is caused by the lack of TE. This also means that reduced TE in agricultural production results in reduced agricultural production.

2.4.2.2. Technical efficiency

The results showed that TE of agriculture in Vietnam is in the range of 21.40% to 98.37%. The average TE of the agricultural sector in Vietnam is 89.29% (Table 2.5a). It shows that 10.71% is lost due to: (i) inefficiencies in agricultural production (either crop or livestock or both) and (ii) inefficiencies among households. Concurrently, results indicate that farmers can increase production by about 9.23% [that is, 1 - (89.29 / 98.37)] by improving in technical efficiency. Looking at TE for each of the province, the average TE of agricultural production reachs the highest in Lao Cai (90.51%), followed by Lai Chau (89.09%) and Phu Tho is in the bottom (87.08%). However, the average TE does not vary significantly across the provinces.

Table 2.5a. Technical efficiency by province

Province	N	Mean	Std.Dev	Min	Max
Phu Tho	560	0.8707954	0.1137362	0.214037	0.9825931
Lao Cai	1150	0.9050505	0.0659928	0.4703018	0.9837003
Lai Chau	725	0.8909489	0.0617432	0.526034	0.9834096
Total	2435	0.8929739	0.0796493	0.214037	0.9837003

(Source: Vietnam Access to Resources Household Survey)

Table 2.5b. Technical efficiency by years

Year	N	Mean	Std.Dev	Min	Max
2008	487	0.9401677	0.0411721	0.6025393	0.9837003
2010	487	0.9221306	0.0521087	0.5121671	0.9785283
2012	487	0.8990930	0.0651741	0.413238	0.9717391
2014	487	0.8699506	0.0803373	0.3112358	0.9628441
2016	487	0.8335275	0.0972963	0.214037	0.9512207
Total	2435	0.8929739	0.0796493	0.214037	0.9837003

(Source: Vietnam Access to Resources Household Survey)

Table 2.5b presents a summary of the average TEs over the study period. The average TE decreased over the years. In 2008, it was 94.02%, 92.21% in 2010, 89.91% in 2012, 87.00% in 2014 and 83.35% in 2016. It can be explained by the following reasons: (i) The process of industrialization and modernization of the country has led to a restructuring of industries. The proportion of industry and trade tends to increase while the proportion of agriculture tends to decrease and lessen people's interest in agricultural production; (ii) Input factors in agricultural production still have positive impact on agricultural output, but the opportunity cost of agricultural production is smaller than the opportunity cost of other industries. This has led to input factors such as labour and other costs of the farmer households gradually shifting to other occupations; (iii) Some typical rural households have a strong transformation in agricultural production and achieved high success in agricultural production while other farmers have maintained the same old production methods. This creates a widening gap between rural households in the study areas.

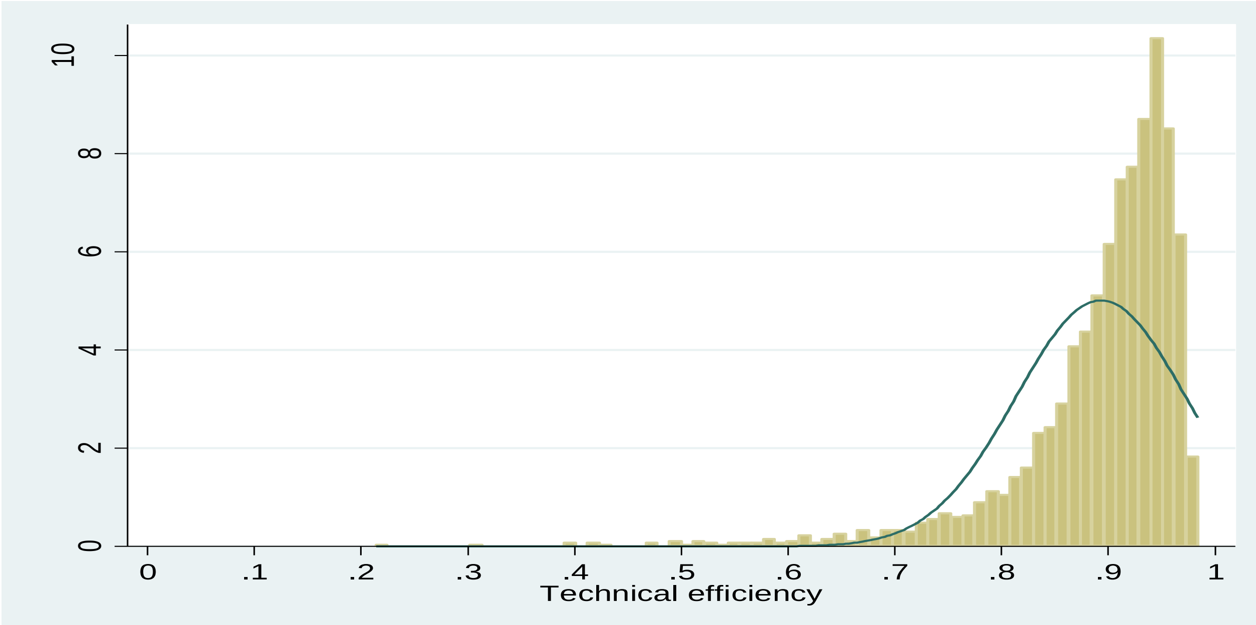


Figure 2.5. Distribution of technical efficiency in Vietnam agriculture

Figure 2.5 gives a picture of the distribution of TE in Vietnam agriculture. In which 1≥TE≥0.9 accounted for 60.16%, 0.9> TE≥0.8 accounted for 30.10%, 0.8> TE≥0.7 accounted for 6.41%, 0.7> TE≥0.6 accounted for 2.05%, 0.6> TE≥0.5 accounted for 0.78% and 0.9> TE accounting for only 0.49%. Thus, the allocation density of TE in the range greater than 0.8 and less than 1 accounts for a large proportion (over 90%).

Table 2.6. Factors affecting on technical efficiency

Variables	Coef	Std. Err.	[95% Conf. Interval]	
gender	0.0006184	0.0066343	-0.01123946	0.0136214
family_mem	0.0022846***	0.0008306	0.0006567	0.0039126
ethnic	-0.0348273***	0.0070700	-0.0486842	-0.0209705
land_frag	0.0038163***	0.0008307	0.0021882	0.0054444
Constant	0.8635241***	0.0104193	0.8431027	0.8839456

(Source: Vietnam Access to Resources Household Survey)

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

We used Wald test to test whether the model covariates significantly influenced TE and found that gender of household head, number of family members, ethnicity of household head, number of plot of the household improved the model fit.

The coefficient of family_mem is positive, indicating that households with more family workers realize higher TE than family labour-constrained households. In Vietnam, farming systems are labour intensive and family labour is critical in the context of resource poor farmer to increase TE in production. The coefficient of ethnic variable is negative, explaining that the Kinh-headed or Hoa-headed households have lower TE than other ethnicities. This is entirely appropriate because the Kinh and Hoa groups often allocate to live in delta while the study area is mountainous and the Kinh ethnic group accounts for a small proportion and less advantage than the ethnic minority peoples. The coefficient of land_frag is positive, indicating that households with more plots have an advantage over households with less plots. In other words, households with more plots have higher TE than households with less plots. This is very true in Vietnam, especially in the northern mountains where there is strong fragmentation of land. The large number of plots often means that the land area is larger, and the larger the number of plots, the easier it is to allocate the cultivation and husbandry production structure than households with less land.

2.4.2.3. Elasticity output and return to scale

The results of the elasticity estimation are shown in Table 2.7. The estimated results show that all three inputs affected TE in agricultural production. The most important factor was land area, followed by intermediate costs and then the number of laborers. Specifically, if other inputs were kept constant, as 1% increase in cultivated land, labour or intermediate costs increased the value of production by 33.76%, 24.69%, 20.03%, respectively.

Table 2.7. Estimate input distance elasticity

Input	Elasticity	Std.dev	Min	Max
elland	0.3376601	0.0123452	-0.4048224	0.7143449
ellabor	0.2003111	0.0169819	-0.1545601	0.9417963
elinter	0.2469162	0.0107567	-0.1190930	0.3773170

The RTS estimates are shown in Table 2.8. The average RTS for the period 2008-1016 was 0.7848. It means if all inputs were increased by 1%, the value of agricultural output would increase by 78.48%. RTS over the years increased significantly. This indicates that farmers in northern Vietnam were adjusting agriculture in the right direction by adjusting and combining inputs as well as making a reasonable decision in balancing crop and livestock production. Specifically, when all input factors increased by 1%, the value of agricultural output increased by 71.25% (2008), 73.67% (2010), 80.06% (2012), 81.96% (2014), 85.00% (2016).

Table 2.8. Return to scale by years

Year	N	Mean	Std.Dev	Min	Max
2008	487	0.7125864	0.0890916	0.2891314	0.941106
2010	487	0.7366667	0.1217349	0.291645	1.046581
2012	487	0.80055944	0.1039568	0.3541342	1.074948
2014	487	0.8195803	0.1035733	0.4795656	1.080677
2016	487	0.8500093	0.1046214	0.3701053	1.132897
Total	2435	0.7848874	0.1171131	0.28911314	1.132897

(Source: Vietnam Access to Resources Household Survey)

2.4.2.4. Technical change, total factor productivity and its components

We now compute the components of TFP and TFP. Based on formula (5a) we calculate scale effect, technical inefficiency change, and technical change or frontier shift. However, when calculating technical change or frontier shift we will skip the first year in the calculation to suit the other components to be calculated.

By estimating the average TC value in stochastic frontier production, it is said that in the whole study period, the average ratio of TC in each period is -4.43%. This research result is higher than that of Ho (2012) in the period 1990-2006 when using SFA function to estimate. Specifically, the shift of the production frontier trended down but with uneven proportions. The year with the highest decrease was 2010 (6.42%), then by 2014 (4.37%) and the lowest was 2016 (2.97%). (Table 2.7). The shift inward shift in production frontier may be explained by: (i) The expansion of fiscal policy through a strong increase in money supply that pushed the inflation rate in Vietnam between 2008 and 2012 (specifically, CPI changed from 22.97% to 9.21%) while the prices of agricultural products in general have been changed unevenly with the increase rate of inflation. This leads to the output value calculated by uneven price with output over time periods; (ii) The opportunity cost of the agricultural production sector is lower than that of other sectors, so there is a shift in land and labour from the agricultural to other sectors.

The TFP growth calculation results are presented in the last column of Table 2.9. In general, TFP growth grew in almost every period, except in 2012, with an average annual rate of 2.71%. However, the estimation of TFP growth also showed a clear downward trend. In 2010, TFP growth reached an impressive figure (14.9%), by 2014 it decreased to 3.38%, in 2016 decreased to 1.35%, even in 2012 it reached a negative figure of 8.77%.

Analysis of TFP growth also identifies factors that motivate and inhibit it. In particular, the estimation of technical change (TC= $-\delta \mathbf{lnD_{it}^{0}}/\delta \mathbf{t}$) in the third column of the table has a negative impact on TFP growth during the whole research period. On average, during the whole research period, it contributed to a large decrease (-163.48%) of the total components of TFP growth. The remaining two factors, technical efficiency change and scale effect, mostly have positive effects on TFP growth but are uneven over the period. This shows that the shift inward shift in production frontier has a large and negative impact on TFP growth.

Table 2.9. Productivity change and its components

Year	**N**	**TC**	**TEC**	**SC**	**TFP**
2008	487	.	.	.	.
2010	487	-0.0642126	0.0230613	0.1901846	0.1490333
2012	487	-0.039609	0.0304587	-0.0785517	-0.087702
2014	487	-0.0436856	0.0402288	0.0372955	0.0338387
2016	487	-0.0296685	0.0531329	-0.0099509	0.0135135
Total	2435	-0.0442939	0.0367240	0.0346645	0.0270946

(Source: Vietnam Access to Resources Household Survey)

2.5. Conclusion and policy implications

We have innovatively employed distance function approach on farm household panel data collected between 2008 and 2016 to analyze TE, technical change and return to scale in Vietnam agriculture, as well as Tobit estimator to analyze the drivers of TE in crop and livestock production. The approach is appropriate in scenarios of multiple outputs and inputs and imperfect information about targets such as in Vietnam agriculture. After converting variables that were measured in money into the fixed-price of 2008 through CPI index, we then estimate the parameters of the translog output function using MLE method.

The model results showed a stable level of TE at 89.29% over the study period, but with variation across Phu Tho, Lao Cai and Lai Chau provinces. From output elasticity analyses, we found that all the selected inputs (covariates) positively and significantly influenced the value of agricultural production, in which land was ranked first, followed by intermediate cost and lastly, labour costs. The technical change rate for the whole period was -4.43% and tended to increase over the years, an indication of negative shift of the output distance function inwards. However, the rate of change decreased over the years. The return to scale estimates indicated that the return to scale increased over the study period and its average was 78.48% over the period 2008-2016. Average TFP growth during the period reached 2.71%, however, there was a downward trend over the years.

The study findings have policy implications on TE and return to scale in crop and livestock production in Vietnam as follows. From agricultural production perspective, increasing TE would require scaling up investment in agricultural extension services to increase capacities of farm households in terms of adoption of productivity-increasing technologies that also improve labour efficiency and expanding their market access. Looking at the crop production, adoption of some conservation agriculture practices such as crop rotation, intercropping and multi-cropping whenever suitable can improve soils in the context of intensive agriculture as in Vietnam. Farmers need to optimize selection of inputs used in crop and livestock enterprises for better yields and incomes. In livestock, it is necessary to increase the number of cycle in the year.

From off-farm sector perspective, farm households are less diversified out of or with agriculture into other sectors and the indicators of economic efficiency of land, labour and intermediate costs in agricultural production in Vietnam are quite low. Especially for labour, the revenue or profit per labour of the group with high TE is higher than the revenue or profit on labour of the group with low TE. However, the efficiency of labour in the agriculture is still low. In addition, farmers have a lot of idle time. Therefore, farm households need to diversify their income sources by involvement in off-farm activities such as off-farm labour employment (domestic or foreign), home craft, trading, and services. To ensure this, Vietnam Government should invest in improving skills of the workforce through on-job and vocational training to expand employment opportunities and also ease disguised employment in the agriculture sector. One of the issues that the government needs to solve is macroeconomic management through fiscal and monetary policies to reduce the impact of inflation on agriculture.

3. Modeling technical efficiency using stochastic frontier production function for panel data: An application in rice farming in Vietnam

Abstract

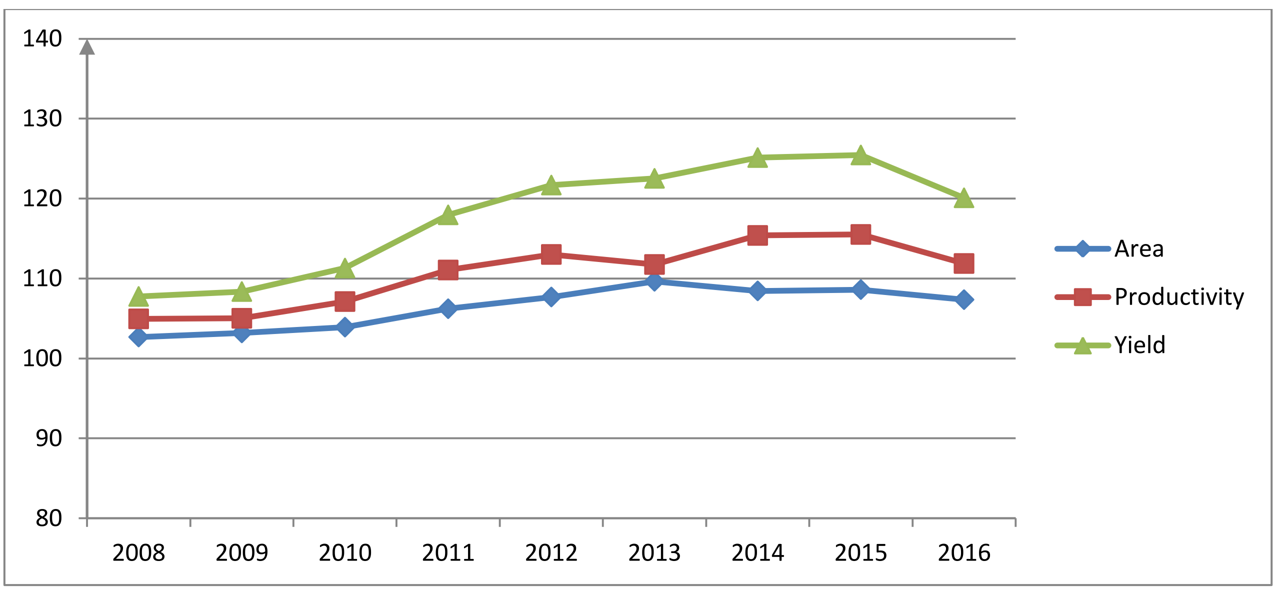
This study analyzes rice farming households in the period 2008-2016. The study focuses on the level of technical inefficiency in the whole Vietnam as well as identifying the factors affecting the technical efficiency. The stochastic frontier with different assumptions is used to establish 6 models applied on a balanced panel data (1555 households, 5 surveys, 7775 observations). By using the likelihood ratio test, the most optimal model was selected. Afterward, tobit model was used to identify factors that influence technology efficiency. The results show that TE score of Vietnam is 92.62% and increased over the years. There are TE differences among the six economic regions, the highest is the North Central Coast, followed by the Red River Delta and the lowest is the Central Highland. The results on output elasticity indicates that all inputs positively influence the quantity of rice production. In which, area cultivated has the strongest influence, followed by fertilizer, seeding, labour, other cost, pesticide and herbicide. Return to scale is 0.93 and increases during the study period. The average growth of TFP was 4.29%. However, the difference in TFP between years is not too large. The results also show that the gender, education and highest level of the household leader, irrigation index and fragmentation positively influence technical efficiency while ethnic of household leader negatively influences technical efficiency. Various suggestions to improve efficiency, technical change and returns to scale are discussed.

3.1. Introduction

Rice is very suitable for humid environments and is grown widely throughout Asia. It is the main food for more than half of the world's population (Khush, 2005). The rapid increase in population has put great pressure on global food systems to feed the growing population. This has also led to a decline in arable lands and increased intensification of production amid declining productivity trends in the face of climate change and weather variability, that adversely affect efficiency in agricultural production. Increasing productivity and efficiency in production is therefore critical particularly in the production of staple food crops. Policy makers are thus placing much emphasis on improving technical efficiency (TE) in rice production.

After the reunification of the country in 1975, Vietnam faced many challenges and difficulties, one of which was a centrally planned economy that resulted in low economic growth and increased poverty. There was pressure to reduce poverty and eliminate hunger while providing food to the rising population. Consequently, Vietnam has built on the ‘Doi Moi’ reforms of 1986 to transform the economy from a centrally planned one to an open market economy with the State's management under the socialist orientation. The reform significantly increased agricultural productivity and Vietnam has transformed from a net importer to net exporter of key agricultural commodities such as rice, cafe and cashew nuts. The country is the second largest exporter of rice.

Rice is the main food of Vietnamese people and is widely grown in Vietnam under different agro-ecological typologies. The country has a long history of wet rice agriculture and paddy is the most important crop. In 2016, rice production area reached 7207.4 thousand hectares, productivity was 4.99 tons/ha and a yield of 35942.7 thousand tons (GSO, 2017). In general, the growth rate in area cultivated, productivity and output of rice have been remarkable since 2007. However, there was a downward trend in 2016 (Figure 1). With 9.3 million farming rice, its production contributes 40% of total agricultural output (Ha, 2012).



(Source: General Statistics Office of Vietnam 2017)

Figure 3.1. Growth rate in area, productivity and yield in the period of 2008-2016 (2007 = 100%)

There have been several studies on TE of rice production in Vietnam. These studies often estimate TE levels using stochastic frontier analysis (SFA), data envelopment analysis (DEA) or both. Previous studies have focused on specific areas such as Mekong River Delta (Hien, Kawaguchi, & Suzuki, 2003), Dong Thap province (Le, Pai Po Lee, Ke Chung Peng, & Chung, 2017), Central Vietnam (Pedroso et al., 2018), Kien Giang province (Tuan M.Cao et al., 2017). Khai and Yabe (2011) conducted a study of TE analysis of rice production across Vietnam, using a Cobb-Douglass (CD) stochastic frontier model to measure TE levels from Vietnam Household Living Standards Survey (VHLSS) data in 2006. However, this study did not provide conclusive evidence of the impact of regional variations on rice production results. No study has analyzed TE in rice production across regions of Vietnam so far. Therefore, extensive regional research to provide evidence of TE levels between regions across Vietnam is needed to answer questions for the planning of rice-production to meet the domestic demand for rice and for the export market.

Agricultural production including rice production suffers from many production and marketing risks. Therefore, short period (i.e. 1 year) analysis of TE in rice production does not give a true picture of efficiency in production given the agro-ecological heterogeneity in Vietnam and data limitations in analyzing trends in technical change as observed in many cross-sectional studies (Hien et al., 2003; Hoang Linh, 2012; Khai & Yabe, 2011; Le et al., 2017; Pedroso et al., 2018; Pham, 2016; Tuan M.Cao et al., 2017).

The selection of a suitable model in terms of structure and distribution assumptions of technical inefficiency is important in predicting TE. Some studies have used CD or Translog as default function of stochastic frontier without verification (Hien et al., 2003; Hoang Linh, 2012; Huynh-Truong, 2009; Khai & Yabe, 2011; Pedroso et al., 2018; Tuan M.Cao et al., 2017). Most of these studies assumed either truncated normal U_i distribution (Hien et al., 2003; Kompas, 2002; Pedroso et al., 2018) or half-normal distribution (Hoang Linh, 2012; Khai & Yabe, 2011), and other modeled it as unknown (Huynh-Truong, 2009; Pham, 2016; Tuan M.Cao et al., 2017), which may result in deviations in calculating TE.

When examining the drivers of TE, many studies only focus on the internal factors in the rice farming households and ignore the external factors such as irrigation rates, disasters as well as institution drivers such as land fragmentation and existing agricultural and trade policies (Hien et al., 2003; Hoang Linh, 2012; Huynh-Truong, 2009; Pham, 2016). Land fragmentation often negatively influence crop yields and increases other production costs (Van Hung, 2007). The central issue in the context of Vietnam is to understand which factors influence TE, the direction of the effects and mechanisms to adjusts the factors to yield optimal outcomes.

In this paper, we contribute to the knowledge gap on TE in rice production in Vietnam in four ways. Firstly, unlike other related studies that used cross-sectional data, this is the first experimental study of TE of rice production in 6 economic regions of Vietnam using a balanced panel data of 1555 farm households in each of the five survey rounds (7775 observations). Secondly, employ various assumptions about the production of stochastic frontier, the distribution of U_i to find the most appropriate production stochastic frontier for the rice production data. Thirdly, on the basis of the most relevant production stochastic frontier, the study ranks TE of rice farmers by provinces (regions) and years. Lastly, we consider internal and external factors in analyzing the drivers of TE to provide a holistic view of the production environment and mechanisms to increase TE.

3.2. Literature review

3.2.1. Theoretical background of stochastic frontier analysis

SFA takes its source from Meeusen and van den Broeck (1977) and Aigner, Lovell, and Schmidt (1977). Both articles have a few similarities. Shortly thereafter appeared a third paper written by Battese and Corra (1977). Model is expressed as:

$$Y = f(X;\beta).\exp(V-U) \qquad (1)$$

Where Y is scalar output, X is a vector of inputs, β is a vector of technology parameters, the symmetric random error V accounts for statistical noise and production risk. All three articles focus on analyzing the structural error. The first error in the structure is $V \sim N(0, \sigma_v^2)$ to represent the effect of statistical noise. The second error in the structure is $U_i \geq 0$ for the effects of TE. The difference between them is the believing on the distribution of U_i, the distribution of U_i conformed to an exponential distribution (Meeusen & van den Broeck, 1977). It was a half-normal distribution (Battese & Corra, 1977). Meanwhile, it could have both types (Aigner, Lovell, & Schmidt, 1977). A few years later, Green (1980a, b) discovered Gamma distribution and Stevenson (1980) formulated Gamma and truncated normal distribution. Schmidt&Lovell (1979) completed the CD function.

The convenience of the SFA model can be explained by graphs. For simplicity, we assume 1 input X_i to produce output Y_i. Figure 1 shows the production frontier for two companies, A and B. Firm A uses input level X_A to produce output Y_A and firm B uses input level X_B to produce output Y_B. In the case of firm A, it shows very clearly that the frontier output for firm A lies above the deterministic frontier because the noise effect is positive ($V_A > 0$), while the fuzzy output for firm B lies below the deterministic frontier because noise effect is negative ($V_B < 0$). We can assume that the stochastic frontier output varies from the deterministic frontier by noise affect. Both firms can improve their efficiency by changing TE so that outputs are feasible with corresponding inputs.

Most studies take an output-oriented approach to measure TE by the ratio observed output to the corresponding stochastic frontier output. The stochastic frontier output varies from the deterministic frontier by noise affect and the TE value is between 0 and 1.

$$\mathrm{TE} = \frac{\mathrm{Yi}}{\exp\left(X_i'\beta + V_i\right)} = \frac{\exp\left(X_i'\beta + V_i - U_i\right)}{\exp\left(X_i'\beta + V_i\right)} = \exp(-\mathrm{Ui}) \qquad (2)$$

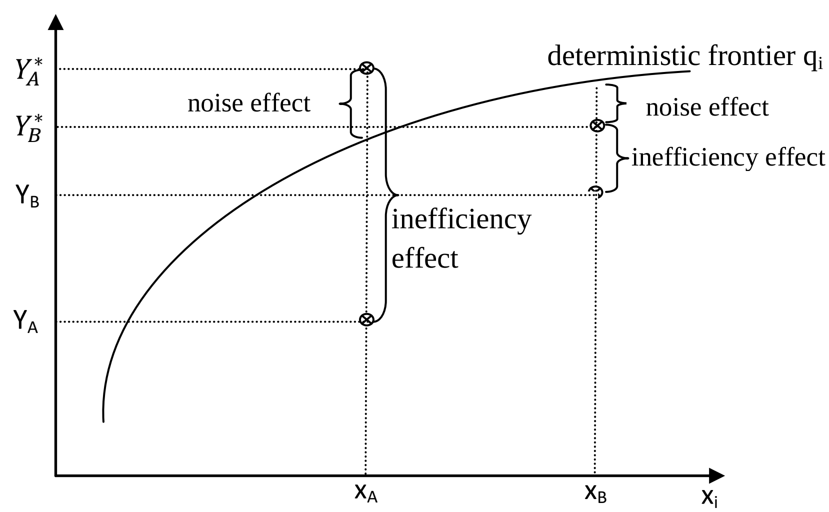


Figure 3.2. The stochastic frontier model (Battese & Coelli, 1995)

Cross-sectional data only provides a partial picture of the efficiency of producers. The panel data gives us a more general and detailed picture of their transition over time periods. A panel data specification is formulated as follows:

$$Y_{it} = f(X_{it}; \beta)e^{Vit - Uit} \qquad \mathrm{Uit} > 0;\ i = 1 \ldots N;\ t = 1 \ldots T \qquad (3)$$

There are 2 types of relationship between technical inefficiency and time, they are time-invariant and time-variant model. Time-invariant model assumes that the technical inefficiency is fixed over time (U_i) (Battese & Coelli, 1988; Schmidt & Sickles, 1984) while time-variant model allows technical inefficiency to change over time (U_{it}) (Battese & Coelli, 1992; Cornwell et al., 1990; Kumbhakar, 1990). Schmidt and Sickles (1984) applied fixed effects and random effects methods to measure efficiency with the assumption of time-variant efficiency as Battese and Coelli (1988) assumed time-invariant followed $U_i \sim N(\mu,\sigma^2)$ distribution.

3.2.1.1 Time-invariant models

**** Fixed effect model - Schmidt and Sickles (1984)***

In the fixed effect model, there are no hypothetical distributions on U_i and correlate with regression variables as well as V_{it}. Schmidt and Sickles (1984) assume that f (.) is lined in the x (i.e log of inputs and outputs in the Cobb-Douglas production function. Thus, the fixed effects models can be written as follows:

$$Y_{it} = \beta_0 + X'_{it}\,\beta + V_{it} - U_i \tag{4}$$

$$Y_{it} = \alpha_i + X'_{it}\,\beta + V_{it} \qquad \text{where } \alpha_i = (\beta_0 - U_i) \tag{5}$$

In this model, α_i and u_i are assumed as a fixed parameter estimated together with the parameter vector β. Thus, this model can be estimated by OLS.

According to Schmidt and Sickles (1984)

$$\widehat{U_i} = \max(\hat{\alpha}) - \hat{\alpha}_i \geq 0, \quad \text{where } \hat{\alpha} = \max(\hat{\alpha}_i) \tag{6}$$

This formula implies the assumption that the highest effectiveness unit of the sample is 100% effectiveness. Consequently, inefficiency estimation in the fixed effect model is the comparation with the best unit in the sample.

**** The random effects model - Schmidt and Sickles (1984)***

In the random effects model, the Ui is assumed to be randomly distributed with constant mean and variance and is not correlated with regression variances as well as V_{it}. At this time, the random effects model provides more efficient estimates than the fixed effects model.

$$Y_{it} = \beta_0 + X'_{it}\,\beta + V_{it} - U_i \tag{7}$$

let $E(U_i) = \mu$ and $U^*_i = U_i - \mu$ and $\alpha^* = (\beta_0 - \mu)$

$$Y_{it} = (\beta_0 - \mu) + X'_{it}\beta + V_{it} - U_i^* \qquad (8)$$

$$Y_{it} = \alpha^* + X'_{it}\beta + V_{it} - U_i^* \qquad (9)$$

In this case, it to be estimated by the generalized least square (GLS). GLS approach provides $\hat{\beta}$ and $\hat{\alpha}^*$. Then, U_i^* is estimated from either the residuals or the best linear unbiased predictor (BLUP)

$$\hat{U}_i^* = \frac{1}{T}\sum(Y_{it} - \hat{\alpha}^* - \hat{\beta}X_{it}) \qquad (10)$$

$$\text{then } \hat{U}_i = \max\{\hat{U}_i^*\} - \hat{U}_i^* \qquad (11)$$

**** The Battese and Coelli (1988) model***

Battese and Coelli proposed frontier production function

$$Y_{it} = X_{it}\beta + E_{it} \qquad \text{where } E_{it} = V_{it} - U_i \qquad (12)$$

V_{it} is a random error and is defined as independent and identically distributed $N(0, \sigma^2{}_v)$; U_{it} is assumed to be independent and identically distributed non-negative truncations of the $N(\mu,\sigma^2)$ distribution; The technical efficiency of the i*th* firm at the t*th* time period is defined as: $TE_{it} = \exp(-U_i)$

3.2.1.2. Time-varying model

**** The Cornwell, Schmidt, and Sickles (1990) model***

Cornwell, Schmidt, and Sickles (1990) used the form of a fixed effect model with a time-invariant assumption.

$$Y_{it} = \alpha_i + X'_{it}\beta + V_{it} \qquad \text{where } \alpha_i = \beta_0 - U_i \qquad (13)$$

Then, replaced α_i by α_{it}, where $\alpha_{it} = \alpha_{i0} + \alpha_{i1}t + \alpha_{i2}t^2$

Where the parameters $\alpha_{i0}, \alpha_{i1},$ and α_{i2} are firm-specific and t is the time trend variable. Cornwell, Schmidt, and Sickles (1990) define N firm dummies and interaction of these dummies with time and time squared.

$$Y_{it} = \alpha_{i0} + X'_{it}\,\beta + V'_{it} \quad \text{where } V'_{it} = V_{it} + \alpha_{i1}t + \alpha_{i2}t^2 \qquad (14)$$

The form of the model is the same as a panel data model and is similar to that of Schmidt and Sickled (1984). Ordinary Least Squares (OLS) regression was used to estimate $\hat{\beta}$ and the residuals of the model $\hat{\alpha}_{it} = Y_{it} - \hat{\beta} X_{it}$

Then $\hat{U}_{it} = \hat{\alpha}_t - \hat{\alpha}_{it}$ and $\hat{\alpha}_t = \max_j \hat{\alpha}_{jt}$

**** The Kumbhakar (1990) model***

$$Y_{it} = f(X_{it}, \beta) + \varepsilon_{it} \qquad (15)$$

where: $\varepsilon_{it} = V_{it} - U_{it}$; $U_{it} = G(t)U_i$; $V_{it} \sim N(0, \sigma_v^2)$; $U_i \sim N^+(\mu, \sigma_u^2)$

In this model, G(t) is a function of time (t), U_{it} changes over the time and on individuals. U_{it} is composed of non-stochastic time (G (t)) and a stochastic individual (U_i)

Because $U_i > 0$ and $U_{it} > 0$ therefore $G(t) > 0$

The Kumbhakar (1990) model assumed as:

$$G(t) = [1 + \exp(\gamma_1 t + \gamma_2 t^2)]^{-1} \qquad (16)$$

G(t) can be monotonically increasing (decreasing) or concave (convex) depending on the signs and magnitudes of $\gamma_1$and γ_2. The random and nonlinear nature of the model requires iterative estimation by the ML estimation method.

**** The Battese and Coelli 1992 model***

Battese and Coelli proposed a stochastic frontier function with a simple exponential specification of time varying form effects như sau:

$$G(t) = [1 + \exp(\gamma_1 t + \gamma_2 t^2)]^{-1} \qquad (17)$$

$$\text{and } U_{it} = \eta_{it}U_i = U_{it} = \{\exp[-\eta(t - T)]\} \qquad (18)$$

In which:

V_{it} is random errors and is assumed as independent and identically distributed $N(0, \sigma^2_v)$; U_{it} is assumed as independent and identically distributed non-negative truncations of the $N(\mu,\sigma^2)$ distribution; η is an unknown scalar parameter; the technical efficiency of the i*th* firm at the t*th* time period is calculated as follows: $TE_{it} = \exp(-U_{it})$

**** Battese and Coelli 1995 model***

Existing studies have analyzed the drivers (predictors) of TE using a two-stage approach where the first stage estimates the stochastic production function and predict TE, and the second stage implements TE regression and predicts explanatory variables. Two-stage regression is thought to be inconsistent in assumptions regarding the independence of ineffective effects in the two estimated periods. Battese and Coelli (1995) widened their model in the 1992 model. In version 1995, U_{it} is non-negative random variables. It is thought to be technical ineffective in production and independently distributed as truncations at zero of the $N^{+}(\mu_{it},\sigma^2_u)$ distribution. U_{it} is defined as a function of the set of explanatory variables as follows:

$$U_{it} = Z_{it}\delta + W_{it} \tag{19}$$

Z_{it} is a $p \times 1$ vector of variables (such as covariates or time variables) which may influence the efficiency, δ is an $1 \times p$ vector of parameters to be estimated, and W_{it} is defined by a truncation of the normal distribution with zero mean and variance.

MLE is applied for simultaneous estimation of parameters of the stochastic frontier production function (17) and the technical inefficiency effects model (19) (Battese & Coelli, 1995).

**** True fixed effect (TFE) and true random effect (TRE)***

True fixed effect model and true random effect model were implemented by William Greene (Willam Greene, 2005; William Greene, 2005). The main argument is that the ineffective component of fixed effect model and random effect model absorbs the cross-unit heterogeneity, which was presented as regression variables in the function but not as inefficient.

The true fixed effects model is shown as follows:

$$Y_{it} = \alpha_i + X'_{it}\,\beta + V_{it} - U_{it} \tag{20}$$

Where α_i is the unit specific intercept intended to capture all time-invariant heterogeneities; $U_{it} \sim iid\ N^{+}(0,\sigma^2_u)$.

The true random effects model is as follows:

$$Y_{it} = \alpha_i + X'_{it}\,\beta + V_{it} - U_{it} + W_i \tag{21}$$

Where W_i is time invariant unit specific random term designed to capture cross-unit invariant heterogeneity; $U_{it} \sim iid\ N^+(0,\sigma^2_u)$.

**** Wang and Ho 2010 model***

Wang and Ho (2010) eliminate incidental parameters by either first differencing or within transformation in the following model specification:

Recall (20): $Y_{it} = \alpha_i + X'_{it}\,\beta + V_{it} - U_{it}$

Where: $V_{it} \sim N(0, \sigma_v^2)$; $U_{it}=h_{it}\,U^*_i$ where: $h_{it}= Z_{it}\delta$; $U^*_i \sim iid\ N^+(\mu,\sigma^2_u)$.

U_{it} is the multiplicative form of inefficiency effects; U_i is the individual-specific effects; h_{it} is the multiplicative form with the individual and time specific effects. U^*_i does not change with time, the within and the first-difference transformations leave this stochastic term intact.

3.2.2 Some studies on technical efficiency of rice farming in Vietnam

Kompas (2002) used SFA method with unbalanced panel data and cross-sectional data from 60 provinces in Vietnam, from 1991 to 1999 (540 observations in total). The log-linear CD function was assigned. In SFA model, the independent variable was rice yield; dependent variables were: stock of capital, labor, material input, time trend. Inefficiency variables included: Average of farm size, percentage of used tractors, soil condition, number of threshing machines, number of tractors. The coefficients of capital, labour, land, and material inputs are 0.17, 0.13, 0.24 and 0.51, respectively. A time trend also tested significant at 1.1 percent annually. The Red River Delta (RRD) and Mekong River Delta (MRD) are two main areas of rice farming in Vietnam with a 11-13% higher TE than other areas. The model results showed that the model covariates significantly explained technical inefficiency. The study argued that lack of credit markets and land fragmentation stifled agricultural mechanization.

Hien (2003) used the SFA method to estimate stochastic frontier production function incorporating a model for technical inefficiency effects (Battese & Coelli, 1995) is applied to field survey data on 120 paddy farmers of the Mekong Delta. Findings showed that TE means values are 86.23%, 79.55%, and 80.24%, respectively, in the winter-spring, spring-summer and summerautumn seasons. In SFA function, the quantity of active phosphate and potassium and expense for hired machine positively influenced output while quantity of seed, active nitrogen and pesticide costs negatively influenced output.

Dummy variables of land size, use of rice variety, adoption of integrated pest management (IPM) and sowing technique and credit availability positively influenced TE. The study also proposed a raft of measures to improve TE such as land policy reforms to facilitate land accumulation (acquisition) and tenure security to promote investment in pre and post-harvest technologies, expanding credit access and upgrading infrastructure to open input and output markets.

Huy-Truong (2009) analyzed the TE of 261 rice producers in the MRD using a combination of both data envelopment analysis (DEA) and SFA models. DEA estimated TE at 76,% under both Constant (CRS) and Variable Returns to Scale (VRS). The average scale efficiency score for these rice producing households was nearly one. The quantity of rice or yields and of the TE were significantly influenced by variables such as the plot size, seed, and hired labour cost. The farming experience and adoption of advanced farming practices took positively influenced TE.

Khai (2011) also carried out a study on TE in rice production in Vietnam by employing a two-stage SFA model on VHLSS-2006 data from 3373 rice producers. In the first stage, stochatic forntier was estimated for CD production function then TE predicted. In the second stage, tobit function was used to determine the drivers of TE. The TE stood at 81.6%. The findings showed that most of the model covariates positively influenced TE.

Hoang Linh (2012) used a combination of DEA and SFA models to analyze the efficiency of rice producers in Vietnam. The study used data from VHLSS - 2004 and randomly selected one part of the sample with 600 farm households. The results showed a mean TE of 0.704 under CRS, 0.765 under VRS for output-oriented DEA and 0.785 under VRS for input-oriented DEA. The SFA model estimated a mean TE of 0.634. Research results also indicated that TE was positively influenced by education level of household head and land size. TE was highest in the RRD.

Pham (2016) examined the influence of collective action effect on productivity and efficiency of rice producers in Vietnam using SFA method on data from 280 specialty rice producers. Among them, 170 farmers were members of the special rice farmers' association and 110 were not members. The average TE was estimated to be 77%, in which households participating in the specialty rice farmers' association achieved an average of 79.4% and 73.5% for non-members. The study proposed that to improve the efficiency of technology, rice-producing households should join the specialty rice farmers' associations or related marketing groups to market rice collectively and avoid the prohibitive transportation and transaction costs involved in individual marketing.

Le (2017) analyzed the drivers of TE of rice farms in Dong Thap province of Vietnam using two-stage DEA on data from 2000 farmers. The estimated results were 0.801, 0.829 and 0.966 for overall technical efficiency, pure technical efficiency and scale efficiency of the rice farms, respectively. Both overall and pure technical efficiency were positively influenced by education level of the farmers but negatively affected by credit access and training.

Pedroso (2017) focused on TE of rice production in Vu Gia, Thu Bon river Basin, Central Vietnam using SFA with simultaneous (one-step) estimation of the parameters of exogenous effects on TE. Results indicated that TE was 81% for central Vietnam. TE of this area was affected by the scale of production, fragmentation of the farm and exposure to salinity intrusion risks.

Tuan (2017) conducted research on TE among farmers that belonged to farmer cooperatives in Kien Giang province, MRD using CD model on data from 276 rice producers from 4 cooperatives. The results showed that farmers in the cooperative achieved 92.4% TE in rice production. Farm size, potassium fertilizer, and labour positively influenced output while seed negatively influences output. The study also showed that experience in rice production, attending training courses positively influenced TE.

3.3. Data and Empirical model

3.3.1. Data

Currently, there are two main data surveys in Vietnam on national scale namely, the Vietnam household living standards survey (VHLSS) implemented by the General Statistics Office with support from the World Bank since 1992 and the Vietnam Access to Resources Household Survey (VARHS), implemented by the Institute of Labour Science and Social Affairs under Ministry of Labour Invalids and Social Affairs and funded by Danish International Development Agency (DANIDA). While VARHS data is a panel of sampled farmers while the VHLSS is a census of all households in urban and rural areas.

VARHS data was first collected in 2002 from a sample of 932 households in 4 provinces. In 2006, the sample was increased to 2324 households in 12 provinces. Round 2008 was 3223 households, round 2010 was 3202 households (in which, 2200 panel households),round 2012 was 3700 households, round 2014 was 3648 households and 3582 households in round 2016. After filtering and connecting data, the panel data used in this study including 1722 rice producing households in the period of 2008-2016 in 12 provinces and 7 economic regions. However, data from the South Central Coast region (Quang Nam and Khanh Hoa province) was excluded in the analysis due to significant errors noted.

This paper uses a 5-rounds balanced panel (VARHS) data from 1555 rice farmers collected every 2 years from 2008 to 2016 in 6 economic regions namely, RRD - Ha Tay province , North East NE - Phu Tho province, North West NW - Dien Bien, Lao Cai, Lai Chau provinces, North Central Coast NCC - Nghe An province, Central High land CH - Dak Lak, Dak Nong, Lam Dong provinces, MRD - Long An province) as listed in table 3.1.

Table 3.1: Observations by regions

Detail	RRD	NE	NW	NCC	CH	MRD
Number of household	240	169	740	101	204	101
Observation	1200	845	3700	505	1020	505

(Source: Vietnam Access to Resources Household Survey)

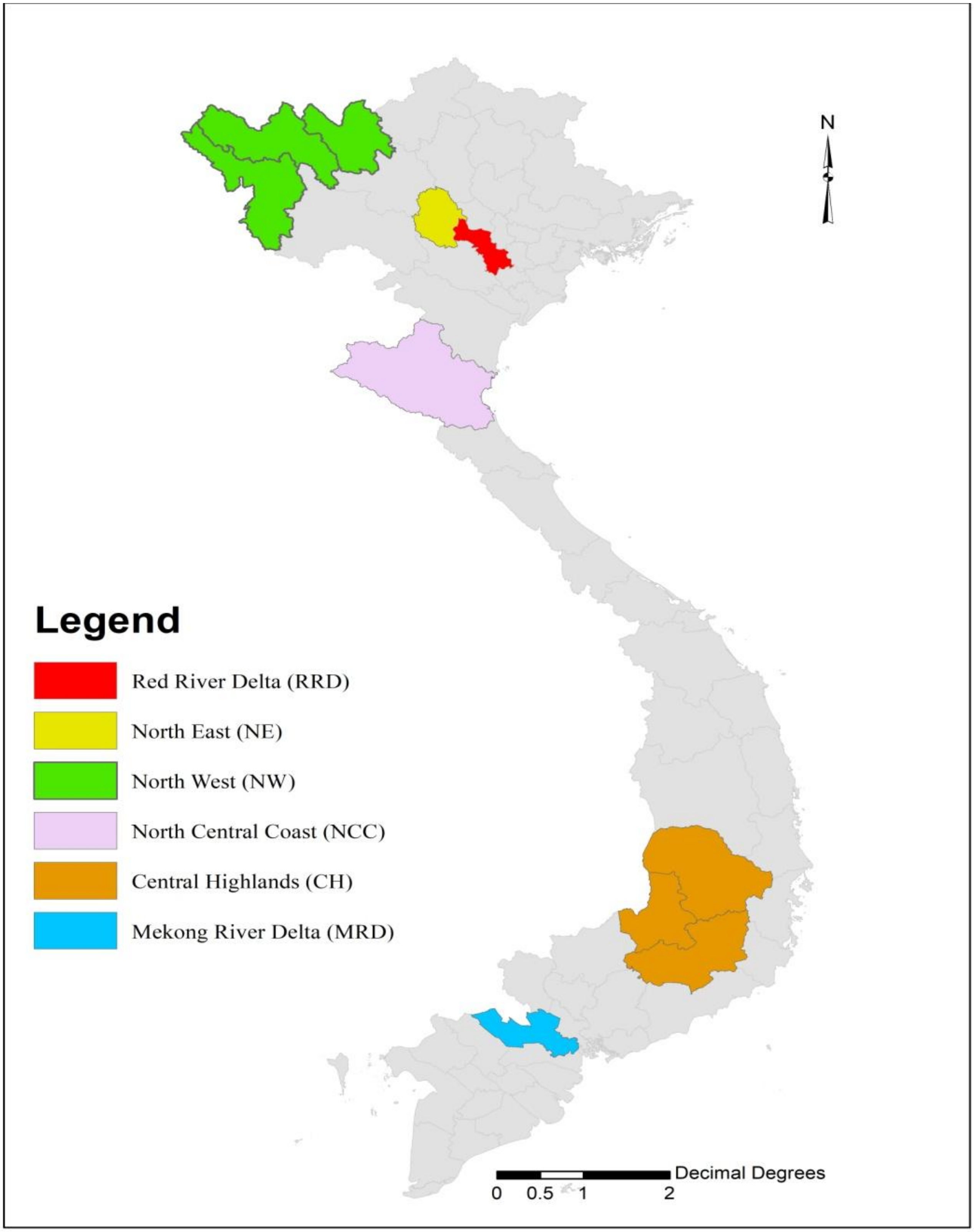


Figure 3.3. Map of research areas

3.3.2. Variables

The study was conducted through the monitoring of all inputs and outputs in the rice production process of farmers in Vietnam. In the past 10 years, inflation has been a problem that has significantly impacted on Vietnam's economy in general and agricultural production. Therefore, for comparisons over time, the quantifiable variables were left intact while those measured in terms of money value were converted to fixed prices in 2008 using consumer price index (CPI) for Vietnam.

Output in rice production includes rice and rice stalks. However, in recent years, rice stalks are usually destroyed instead of being used as a source of fuel and cattle feed. Therefore, in this study, the output is the quantity of rice produced by farmers.

The area of rice cultivation is an important input factor in agricultural production in general and rice production in particular. Thus, it is an indispensable input in stochastic frontier production of rice (Hien et al., 2003; Hoang Linh, 2012; Huynh-Truong, 2009; Khai & Yabe, 2011; Le et al., 2017; Tuan M.Cao et al., 2017). In this study, the area of rice cultivation is strictly counted by the rice cultivation area of each field for each crop in the study year.

Labour is also an important input in production. In rural of Vietnam, the main source of labour is family members. There is also widespread reciprocated labour (labour exchange) between households during peaks of the seasons. Labour is often measured by labour per day and it is an important input in stochastic frontier production of rice (Hoang Linh, 2012; Huynh-Truong, 2009; Khai & Yabe, 2011; Tuan M.Cao et al., 2017).

Seeds are usually bought from farmer cooperatives or retailers. However, in some places, especially in the northern areas, seeds are recycled from the previous harvests. Seed costs were adjusted for inflation.

Fertilizers include chemical and organic fertilizers. Chemical fertilizer is offend bought from cooperatives or retail suppliers. Organic fertilizers can be home-made or buy in the market. Fertilizer costs are converted into the fixed price in 2008. Pesticides, herbicides, and stimulants are bought from cooperatives or retail suppliers. Those costs were also adjusted for inflation. Other costs included the costs of hired labour and draught power, other material inputs such as sickles, shovels, bamboo baskets and fuel.

3.3.3. Empirical estimation

In the first stage, the maximum likelihood estimation of the parameters of the stochastic frontier production function was employed using the Stata frontier function for the different models. There are 2 steps to find the best models: (i) Chose the structure of stochastic frontier production function of models: Log-likelihood ratio (LR) test was used to chose if Cobb-Douglas stochastic frontier production form or translog stochastic frontier production form; (ii) Based on the different assumptions of the existence of Gamma, Mu, Eta combination, LR test was used to find the best model. Based on the best model, output elasticity and scale elasticity were estimated by production frontier function. Simultaneously, TE of rice farmers was estimated by regions and years.

The parameters of the stochastic frontiers is estimated using the Cobb-Douglas production function (CD) as follows:

$$lnY_{it} = \beta_0 + \beta_1 lnX_{1it} + \beta_2 lnX_{2it} + \beta_3 lnX_{3it} + \beta_4 lnX_{4it} + \beta_5 lnX_{5it} + \beta_6 lnX_{6it} + yr + regi + V_{it} – U_{it}$$

In which: ln denotes the natural logarithm of variables; the subscript i represents the ith sample household; the subscript t represents the t*th* year; yr denotes a time trend; V_{it} is i.i.d. N(0, σ2v); Uit i.i.d. N(μ,σ2); Y represents for the quantity of rice yield; X1 represents area of rice cultivation land; X2 represents total number of days spending on rice production; X3 represents seed/seedling expenditure; X4 represents fertilizer expenditure (chemical and organic fertilizers); X5 represents pesticides and herbicides expenditure; X6 represents all other expenses (hiring labour or livestock, renting agricultural machinery, etc.).

Then, a translog function was estimated with the same independent variable and the extended explanatory variables are as follows:

$$lnY_{it} = \beta_0 + \beta_1 lnX_{1it} + \beta_2 lnX_{2it} + \beta_3 lnX_{3it} + \beta_4 lnX_{4it} + \beta_5 lnX_{5it} + \beta_6 lnX_{6it} + \delta 1yr + 0.5\beta_{11} lnX_1\_sqit + 0.5\beta_{22} lnX_2\_sqit + 0.5\beta_{33} lnX_3\_sqit + 0.5\beta_{44} lnX_4\_sqit + 0.5\beta_{55} lnX_5\_sqit + 0.5\beta_{66} lnX_6\_sqit + 0.5\delta_{11} yr\_sq + \beta_{12} lnX_{1it} lnX_{2it} + \beta_{13} lnX_{1it} lnX_{3it} + \beta_{14} lnX_{1it} lnX_{4it} + \beta_{15} lnX_{1it} lnX_{5it} + \beta_{16} lnX_{1it} lnX_{6it} + \delta_{1it} lnX_{1it} yr + \beta_{23} lnX_{2it} lnX_{3it} + \beta_{24} lnX_{2it} lnX_{4it} + \beta_{25} lnX_{2it} lnX_{5it} + \beta_{26} lnX_{2it} lnX_{6it} + \delta_{2it} lnX_{2it} yr + \beta_{34} lnX_{3it} lnX_{4it} + \beta_{35} lnX_{3it} lnX_{5it} + \beta_{36} lnX_{3it} lnX_{6it} + \delta_{3it} lnX_{3it} yr + \beta_{45} lnX_{4it} lnX_{5it} + \beta_{46} lnX_{4it} lnX_{6it} + \delta_{4it} lnX_{4it} yr + \beta_{56} lnX_{5it} lnX_{6it} + \delta_{5it} lnX_{5it} yr + \delta_{6it} lnX_{6it} yr + regi + V_{it} + U_{it}$$

In this study, the Battese and Coelli 1992 model was used to run different models based on different assumptions for analysis as follows. First, model structure is tested: (i) model 1 is translog stochastic frontier production form, the time function is a time-varying decay model and U_{it}

distribution is independent and identically distributed non-negative truncations of the N (μ, $\boldsymbol{\sigma_u^2}$) distribution, there is technical change, (ii) model 2 is Cobb-Douglas production frontier form, time function is a time-varying decay model and U_{it} distribution is independent and identically distributed non-negative truncations of the N (μ, $\boldsymbol{\sigma_u^2}$) distribution; (iii) model 3 is the same with model 1 but without technical change. After that, the best model is chosen.

Secondly, all components of the best model are tested. (i) model 4 assumes that Gamma = 0, inefficiency component is absent in the model; (ii) model 5 assumes that Mu = 0, this is a time-varying decay model, and U_{it} distribution is independent and identically distributed non-negative and has a half-normal distribution (U_i~iid N(0,$\boldsymbol{\sigma_u^2}$)); (iii) model 6 assumes that Eta=0, this is a time-invariant and Uit distribution is independent and identically distributed non-negative truncations of the N (μ, $\boldsymbol{\sigma_u^2}$) distribution; (iv) model 7 assumes that Mu = Eta = 0, this is a time-invariant and U_{it} is independent and identically distributed non-negative and has a half-normal distribution N(0,σ2) distribution; (v) model 8 assumes that Gamma = Mu = Eta = 0, this is a time-invariant model, and U_{it} distribution is independent and identically distributed non-negative and has a half-normal distribution (Ui~iid N(0,$\boldsymbol{\sigma_u^2}$)). Furthermore, Uit is absent from the model (the observation variables are full technically efficient)

In the second stage, Tobit model will be used to determine the factors that affect on TE with the dependent variable is the TE estimated in step one.

The tobit function is given by:

$$TE = \delta_0 + \delta_1 Z_1 + \delta_2 Z_2 + \delta_3 Z_3 + \delta_4 Z_4 + \delta_5 Z_5 + \delta_6 Z_6 + \delta_7 Z_7$$

In which, dependent variable, TE is TE that was estimated in step 1; Independent variables are: (i) Z_1 represents gender of household leader (1: male; 2: female); (ii) Z_2 represents education of household leader (1: Never went to school; 2: Primary; 3: Secondary; 4: High education); (iii) Z_3 represents highest degree of household leader (1: No diploma; 2: short-term vocational training; 3: long-term vocational training; 4: professional high school; 5: junior college diploma; 6: Bachelor degree; 7: master degree; 8: PhD); (iv) Z_4 represents household leader ethnic (1: if Kinh or Hoa; 0 otherwise); (v) Z_5 represents the result of dividing the total of irrigated paddy land by total of paddy land (1: irrigation <=0.2; 2: 0.2< irrigation<=0.4; 3: 0.4< irrigation<=0.6; 4: 0.8< irrigation<=0.8;5: 0.8< irrigation<=1); (vi) Z_6 represents the result of dividing the total of paddy land which was effected by disaster in the observation year by total of paddy land (1: disaster <=0.2; 2: 0.2< disaster <=0.4; 3: 0.4< disaster <=0.6; 4: 0.8< disaster <=0.8; 5: 0.8< disaster <=1); (vii) Z_7

represents defragmentation index. d_index = 1 - sum (a_i^2/A^2). In which, a_i is the size of each plot, A is total farm area. 0 ≤ d_index ≤ 1. If d index is approaching to 0, the land defragmentation is not much and vise versa.

3.4. Results and discussion

3.4.1. Descriptive statistics

The means of the independent and dependent variables which were used in the stochastic frontier function are shown in the tables 3.2a and 3.2b. The average quantity of rice of the study area is 12.38 million VND. This value sharply decreased from 2008 to 2010 and 2012. The main reason was the high inflation rate in Vietnam during this period. The average area under rice was relatively stable at nearly 8000 square meters per year. The days of labour decreased over the years from 143 days per household in 2008 to 82.33 days per household by 2016. The average labour days reached 109.61 per household during the study period. Seed and fertilizer costs have been stable in recent years, averaging 680 thousand VND per household and 2205 thousand VND per household relatively. Other costs tend to increase by an average of 1782 thousand VND per household. The main reasons are the increase in costs due to the use of machinery for rice production and hiring labour.

Table 3.2a: Description statistics of some compomemts in rice cultivation

Variable	Unit	2008		2010		2012	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Y	kg	2845.10	5603.40	3079.30	6520.23	3551.34	7936.08
X_1	m^2	7503.88	12614.65	7763.68	13519.36	8021.00	14884.20
X_2	Man days	143.17	92.06	115.84	77.04	116.12	87.55
X_3	1000VND	174.99	358.07	778.72	1545.99	779.02	1706.45
X_4	1000VND	2273.77	7371.66	2227.55	5376.36	2091.21	5328.01
X_5	1000VND	888.74	4979.78	862.23	3628.77	674.22	3322.45
X_6	1000VND	1234.84	4122.94	1580.34	4494.91	1991.46	4515.40

(Source: Vietnam access to resources household survey)

(Note: All monetary values have been converted to fixed price in 2008)

Table 3.2b: Description statistics of some components in rice cultivation

Variable	Unit	2014		2016		Average	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Y	kg	3869.18	9499.44	3754.03	10143	3439.79	8130.35
X_1	m^2	7633.69	13865	7706.48	15592.40	7725.75	14131.07
X_2	Man days	90.58	80.08	82.33	71.75	109.61	84.77
X_3	1000VND	868.10	1721.32	802.42	1651.99	680.65	1512.72
X_4	1000VND	2185.27	5456.91	2249.58	6849.65	2205.48	6136.02
X_5	1000VND	688.51	2994.81	766.17	3419.93	775.97	3732.88
X_6	1000VND	1872.89	3463.01	2233.98	6847.12	1782.70	4837.25

(Source: Vietnam access to resources household survey)

(Note: All monetary values have been converted to fixed price in 2008)

3.4.2. Hypotheses tests and parameter estimates

The LR test was used to find the best model. In the first step, find the structure of stochastic frontier production function. Hypothesis test was used to evaluate whether the translog stochastic frontier production form or the CD stochastic frontier production form is fully representative of the rice production and which once is more suitable to the data (H_0: $\beta_{ij} = 0$). The LR test was used to compare model 1 (table 3.3) and model 2 (appendix 3.1). Results provides that the translog stochastic frontier production form (model 1) is more appropriate than CD stochastic frontier production form (model 2). It means the null-hypothesis was strongly rejected. Hypothesis test continues to be used to consider if the appearance of technical change is suitable for data. The LR test to compare model 1 to model 3 (appendix 3.2). The results showed that the model of the technical change was more appropriate for the model without technical change.

After selecting the model structure, based on various assumptions, 5 other models (model 4 to model 8) were built. In the second step, model 1 was taken as the root to compare with other models with different assumptions in order to find the most suitable model. *(i)* In comparison with model 4 (appendix 3.3): the LR test rejected the null hypothesis (H_0: $\gamma = 0$), indicating that the model with the assumption of the existence of U_i is more appropriate than the model with the assumption the absentation of U_i (the observed variables are full technically efficient), thus, model 1 is better than model 4; *(ii)* In comparison with model 5 (appendix 3.4): the LR test rejected the null hypothesis (H_0: $\mu = 0$), therefore the model assuming distribution of U_i was truncated-normal $U_i \sim N(\mu, \sigma_u^2)$) is more appropriate than the model assuming distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), thus model 1 is better than model 5; *(iii)* In comparison with model 6 (appendix 3.5): The LR test rejected the null hypothesis (H_0: $\eta = 0$), indicating that the model assuming is time-varying decay was more suitable than model with the assumption is time-invariant, TE of regions can increase or decrease exponentially depending of the sign of the decay parameter η: when $\eta>0$ it implies that the degree of inefficiency decreases over time and vice versa. In other words, model 1 is better than model 6; (*iv*) In comparison with model 7 (see appendix 3.6), the LR test rejected the null hypothesis (H_0: $\mu = \eta = 0$), indicating that the model assuming the distribution of U_i is truncated-normal $U_i \sim N(\mu, \sigma_u^2)$) and time-varying decay was more appropriate than model assuming distribution of U_i is half-normal distribution ($U_i \sim N(0, \sigma_u^2)$) and time-invariant, thus model 1 is better than model 7; (*v*) In comparison with model 8 (appendix 3.7): The LR test results shows that the null hypothesis (H_0: $\gamma = \mu = \eta = 0$) is rejected, that is to say the model assuming distribution of U_i is truncat-normal $U_i \sim N(\mu, \sigma_u^2)$), time-varying decay with the existence of the U_i is more applicable than a model with the assumption that distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), time-invariant and without the existence of the U_i the model (the observed variables are fully technically efficient), thus model 1 was better than model 8.

Table 3.3. Hypothesis test for model specification and statisstical assumptions

Null Hypothesis	Model	Likelihood Ratio test (LR)	df	Prob > chi2	Decision
H_0: $\beta_{ij} = 0$	model 1 vs model 2	1538.39	28	0.0000	Reject H_0
H_0: $\delta_{1=} \delta_{11=} \delta_{ijt}$	model 1 vs model 3	55.54	8	0.0000	Reject H_0
H_0: $\gamma = 0$	model 1 vs model 4	252.89	1	0.0000	Reject H_0
H_0: $\mu = 0$	model 1 vs model 5	53.46	1	0.0000	Reject H_0
H_0: $\eta = 0$	model 1 vs model 6	102.70	1	0.0000	Reject H_0
H_0: $\mu = \eta = 0$	model 1 vs model 7	123.11	2	0.0000	Reject H_0
H_0: $\gamma = \mu = \eta = 0$	model 1 vs model 8	252.89	3	0.0000	Reject H_0
H_0: $\delta_0 = ... = \delta_7 = 0$	The Wald test	582.77	7	0.0000	Reject H_0

In conclusion, after testing the models by different assumptions, model 1 is considered as the best once. As a result, model 1 is chosen for further analysis.

Table 3.4. Maximum likelihood Estimates of stochastic frontier model (model 1)

Variables (lnY)	Coef	Std. Err.	[95% Conf. Interval]	
lnX_1	0.7589***	0.0617	0.6378	0.8799
lnX_2	0.1977***	0.0567	0.0865	0.3088
lnX_3	-0.0473**	0.0217	-0.0898	-0.0048
lnX_4	-0.0720***	0.0212	-0.1137	-0.0303
lnX_5	0.0623***	0.0237	0.0158	0.1088
lnX_6	-0.0710***	0.0225	-0.1150	-0.0268
yr	-0.0697*	0.0362	-0.1407	0.0013
lnX_1_sq	0.0204**	0.0102	0.0004	0.0405
lnX_2_sq	0.0519***	0.0144	0.0236	0.0802
lnX_3_sq	0.0419***	0.0029	0.0362	0.0476
lnX_4_sq	0.0307***	0.0019	0.0270	0.0343
lnX_5_sq	0.0141***	0.0021	0.0100	0.0183
lnX_6_sq	0.0185***	0.0021	0.0144	0.0226
yr_sq	0.0138**	0.0067	0.0007	0.0269
lnX_1 lnX_2	-0.0460***	0.0095	-0.0646	-0.0274
lnX_1 lnX_3	-0.0165***	0.0029	-0.0222	-0.0109
lnX_1 lnX_4	0.0038	0.0032	-0.0024	0.0100
lnX_1 lnX_5	-0.0097***	0.0037	-0.0169	-0.0025
lnX_1 lnX_6	-0.0038	0.0033	-0.0102	0.0026
lnX_1yr	0.0016	0.0052	-0.0118	0.0087
lnX_2 lnX_3	0.0109***	0.0036	0.0039	0.0180
lnX_2 lnX_4	-0.0081***	0.0037	-0.0153	-0.0008
lnX_2 lnX_5	0.0005	0.0046	-0.0085	0.0095
lnX_2 lnX_6	0.0095**	0.0038	0.0020	0.0170
lnX_2yr	-0.0105*	0.0061	-0.0225	0.0015
lnX_3 lnX_4	0.0021*	0.0011	-0.0000	0.0043

Variables (lnY)	Coef	Std. Err.	[95% Conf. Interval]	
$\ln X_3 \ln X_5$	-0.0001	0.0015	-0.0029	0.0028
$\ln X_3 \ln X_6$	-0.0035***	0.0009	-0.0053	-0.0016
$\ln X_3 yr$	0.0062**	0.0030	0.0003	0.0120
$\ln X_4 \ln X_5$	-0.0018*	0.0010	-0.0037	0.0001
$\ln X_4 \ln X_6$	-0.0015	0.0011	-0.0036	0.0007
$\ln X_4 yr$	-0.0037*	0.0021	-0.0079	0.0005
$\ln X_5 \ln X_6$	-0.0029**	0.0013	-0.0054	-0.0003
$\ln X_5 yr$	0.0030	0.0022	-0.0013	0.0073
$\ln X_6 yr$	0.0097***	0.0022	0.0053	0.0141
2.regi_code	-0.1875***	0.0150	-0.2169	-0.1580
3.regi_code	-0.2276***	0.0129	-0.2529	-0.2023
4.regi_code	-0.2138***	0.0181	-0.2493	-0.1783
5.regi_code	-0.0973***	0.0154	-0.1275	-0.0672
6.regi_code	0.0849***	0.0257	00.0346	0.1352
Constant	0.9708***	0.2335	0.5131	1.44285
Lnsigma2	-1.7949***	0.3274	-2.4366	-1.1531
Ilgtgamma	-0.0911	0.6866	-1.437	1.2547
Mu (μ)	-5.30589	3.4029	-11.7285	1.6107
Eta (η)	0.6386***	0.0839	0.4741	0.8030
$\sigma_S^2 = \sigma_u^2 + \sigma_v^2$	0.1662	0.0544	0.0875	0.3156
Gamma (γ) = σ_u^2/σ_S^2	0.4772	0.1713	0.1920	0.7781
σ_u^2	0.07929	0.0544	-0.0273	0.1859
σ_v^2	0.0869	0.0015	0.0839	0.0898

(Source: Vietnam Access to Resources Household Survey)

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

Results of using maximum-likelihood estimates to estimate the parameters of the stochastic frontier production function with model 1 (table 3.4) shows five input parameters are 99%

confidence interval. In which, the coefficient of lnX_1, lnX_2, lnX_5 have positive and remarkable influence on the output of rice production. It explains that households with large cultivated areas often have an advantage in raising the quantity of rice production. Households who spend a lot of labour means that they are more interested in visiting rice fields, which helps them respond quickly and easily to deal with pests and disasters (risk). Households with large expenditures on pesticides and pesticides will be more proactive in pest control and treatment, thus reducing risks in agricultural production, especially in terms of Vietnam climate, which is favorable environment for development of pests and diseases. This result is similar to that of Kompas (2002), Hien (2003), Khai (2011), Pham (2016), Pedroso (2017), Tuan (2017). In contrast, the coefficient of, lnX_4 and lnX_6 are negative. It explains that output value reduction is due to excessive use of fertilizers and some other unnecessary activities, which is similar to that of Hien (2013) Huy-Truong (2009), Tuan (2017). The coefficient of lnX_3 is negative and statistically significant at 95% confidence interval, which explains that seed cost is not appropriate and may take negatively affecting on the quantity of rice.

Incorporating the variable yr and yr_sq into the production function to measure neutral technical change. Similarly, on the interaction terms between time and other inputs in stochastic frontier function is intended to measure the error rate of technical change. The coefficient of yr and yr_sq were significant at 90% and 95% confidence interval but the coefficient of yr was negative and of yr_sq positive. This indicates that neutral technical change occurred over the period and that technical change grew at an decreasing rate. On the interaction terms between lnX_2 and yr is negative and statistically significant at 90% confidence interval, indicating a technical change in Vietnam's rice production with labour reduces in each period (-1.05%). Similarly, on the interaction terms between lnX_4 and yr is negative and statistically significant at 90% confidence interval, indicating a technical change in Vietnam's rice production, fertilizer costs also decrease at each period (-0.37%). On the other hand, on the interaction terms between lnX_3 and yr is positive and statistically significant at 95% confidence interval, indicating a technical change in Vietnam's rice production, seed costs increase at each period (0.62%). On the interaction terms between lnX_6 and yr is positive and statistically significant at 99% confidence interval, indicating a technical change in Vietnam's rice production with other cost expenditures increase at each period (0.97%). The coefficient estimates on the interaction terms between lnX_1 and yr, lnX_5 and yr are not statistically significant at 90%.

While the time variable in stochastic frontier function captures technical change over time (shifting of the production frontier), in an inefficient equation the time variable captures inefficiency change (changes in the distance of the average unit from the rice production frontier). Indicators Eta (η) in the model is statistically significant at 99%. Eta (η) is positive meaning the technical inefficiency decrease over time, and units move towards the technical frontier (technological catch-up) at a rate of 63.86% per period (two years), in other words, TE increased by the years.

In relation to the region dummy variables parameters, all variables are statistically significant at 99%. By using Red River Delta region as reference for presenting a larger number of observations, it is verifed that the product of some regions are lower in relation to the reference region except Mekong River Delta. This reflects the reality because Mekong River Delta và Red River Delta are the two major rice producing regions in Vietnam, of which MRD is the area which produce the majority of rice export in Vietnam.

In the model, σ_S^2 is estimated at 0.1662, explaining that variance output is caused by technical inefficiencies and random noise. Gamma (γ) is the variance ratio, explaining the total variation in output from the frontier level of output attributed to technical inefficiency. Estimated value of γ is 0.4772. It means 47.72% of total variation in output of rice production is caused by the lack of TE. This showed that if technical inefficiency reduce, total variation in output of rice production can be decreased.

3.4.3. Technical efficiency

The results showed that TE of rice farmers in Vietnam ranged from 16.81% to 99.70%. The average of TE in the whole Vietnam is 92.62%. It can be inferred that 7.38% of the losses are caused by inefficiencies in rice production or inefficiency among households in the sample or a combination of both. Concurrently, results indicate that rice farmers can increase production by about 7.10% [that is, 1 - (90.73 / 98.80)] by improving in technical efficiency. This result is similar to previous research results (Tuan M.Cao et al., 2017).

Table 3.5. Technical efficiency by years

Year	N	Mean	Std.Dev	Min	Max
2008	1555	0.8266946	0.0997045	0.1680629	0.960319
2010	1555	0.9023644	0.0637173	0.3899617	0.9788469
2012	1555	0.9465254	0.0376218	0.6081983	0.9887741
2014	1555	0.9711912	0.0211509	0.7690743	0.9940565
2016	1555	0.9846224	0.0115575	0.8705371	0.9968572
Total	7775	0.9262796	0.0803879	0.1680629	0.9969572

(Source: Vietnam Access to Resources Household Survey)

Table 3.5 presents a summary of the average TEs over the study period. The average TE increased over the years. In 2008, it was 82.67%, 90.24% in 2010, 94.65% in 2012, 97.12% in 2014 and 98.46% in 2016. It can be explained by three aspects (i) The price of rice has little difference between regions, while the rice productivity of households is almost at the maximum level and it is difficult to increase productivity. (ii) Over time, rice-producing households take actions to adjust their inputs as well as incorporate input elements in a reasonable manner in order to optimize production and quantity of output. (iii) Through village living customs, training, experiences sharing, households tend to be similar in the implementation of rice-growing stages such as: seed selection and cultivation skills.

Table 3.6. Technical efficiency by regions

Regions	N	Mean	Std.Dev	Min	Max
Red River Delta	1200	0.9352254	0.0594453	0.5046979	0.9960325
North East	845	0.9335043	0.0639597	0.5226813	0.9944677
North West	3700	0.9256816	0.0823774	0.2444567	0.9965607
North Central Coast	505	0.9392222	0.0620935	0.5105692	0.9937496
Central Highland	1020	0.9115371	0.1037827	0.1680629	0.9968572
Mekong River Delta	505	0.9141490	0.0891457	0.1954592	0.9943066
Total	7775	0.9262796	0.0803879	0.1680629	0.9969572

(Source: Vietnam Access to Resources Household Survey)

The average values of the TE by region are shown in table 3.6. There is significant differences in TE between the six economic zones, this is also explained in the model with the value of γ. In which, the highest average value of TE belongs to North Central Coast (93.92%), followed by Red River Delta (93.52%), North East (93.35%), North West (92.57%), MRD (91.41%), and Central Highland (91.15%). The difference in TE value between regions is negligible. This is perfectly true for rice production in Vietnam, especially in the RRD, where there are favorable natural conditions and least risks of natural disasters and diseases. This result is similar to a previous study (Kompas, 2002). Meanwhile, MRD is a large-scale and have a high level of intensive farming investment. However, in this area, the productivity is low, the quality of rice is not high and there are other risks so TE in rice production in this region is lower than in others.

3.4.4. Output elasticity and return to scale

Due to the combination of inputs in Translog production frontier, parameters in a translog production frontier are not representative for output elasticities. The elasticity of the average quantity of output to the *k*-th input is calculated as follows:

$$\frac{\partial \ \ln E(\text{output})}{\partial \ \ln (X_k)} = \beta_k + 2\beta_{kk}X_{kit} + \sum_{j\neq k}^{5} \beta_{kj}X_{jit}$$

Table 3.7. Estimate output elasticities and return to scale

Input	Elasticity	Std. Dev	Min	Max
LnX_1	0.5886510	0.0496873	0.4146092	0.8134214
$Ln X_2$	0.0766216	0.0434659	-0.1219738	0.2801421
$Ln X_3$	0.0948310	0.0923492	-0.2000892	0.3027880
$Ln X_4$	0.0991932	0.0667721	-0.1092155	0.2584766
$Ln X_5$	0.0304930	0.0224237	-0.0704998	0.0982861
$Ln X_6$	0.0375663	0.0426727	-0.1162563	0.1355516
RTS	0.9273561	0.1250858	0.3398123	1.214847

(Source: Vietnam Access to Resources Household Survey)

The estimation of output elasticity in table 3.7 shows that all inputs in rice production take positive effect on quantity of output. In which, land is the most important input, followed by, fertilizer, seed, labour, other costs and pesticide, herbicide. This means that land expanding impacts significantly on the quantity of output (Shar, Mahesar, Chandio, & Memon, 2017). Specifically, if other inputs are constant, 1% expansion of paddy land leads to 0.59% increase in quantity of output. Other inputs also take positive in yield value but not too much. In details, if all other variables are constant, when increases 1% of labours, seeding, fertilizer, pesticide and herbicide and other costs, yield value increase 0.08%; 0.09%; 0.10%; 0.03%; 0.04% relatively. The scale elasticity can be obtained through the sum of output elasticity and stood at 0.93. It means if all inputs were increased by 1%, quantity of output would increase by 0.93%.

The RTS estimate is shown in Table 3.8. RTS over the years increased significantly. This indicates that rice-producing households in Vietnam have adjusted the right direction through adjusting and combining inputs. Specifically, when all input factors increase by 1%, the quantity of output will increase by 77.48% (2008), 94.83% (2010), 96.24% (2012), 97.65% (2014), 97.46% (2016).

Table 3.8. Return to scale by years

Year	N	Mean	Std.Dev	Min	Max
2008	1555	0.7748616	0.1560582	0.3398123	1.099651
2010	1555	0.9482988	0.0917934	0.5004183	1.178986
2012	1555	0.9624343	0.0791095	0.5386288	1.149181
2014	1555	0.9765130	0.0721877	0.5797026	1.214847
2016	1555	0.9746728	0.0665213	0.4925652	1.161557
Total	7775	0.9273561	0.1250858	0.3398123	1.214847

(Source: Vietnam Access to Resources Household Survey)

On output elasticity and RTS of the translog production frontier, (i) land plays the most important role in increasing the quantity of output; (ii) Overall quantity of output can be improved by investing in more inputs; (iii) RTS increasingly improved during the study period.

3.4.5. Technical change, total factor productivity and its components

Table 3.9. Total factor productivity and its components

Year	N	TC	TEC	SC	TFP
2008	487	.	.	.	.
2010	487	-0.0155383	0.0675262	-0.0033365	0.0456538
2012	487	0.0028099	0.0356562	0.0037760	0.0418385
2014	487	0.0199166	0.0188277	-0.0002525	0.0387670
2016	487	0.0350445	0.0099417	0.0001023	0.0455520
Total	2435	0.0105582	0.0329880	0.0001065	0.0429210

(Source: Vietnam Access to Resources Household Survey)

Việc tính toán Total factor productivity growth dựa vào công thức: TFP = TC + TEC + SC. Trong đó: TC = $\frac{\delta \ln f(.)}{\delta t}$; TEC= $\frac{\delta U}{\delta t}$; SC = (RTS - 1)$\sum_j \frac{\varepsilon_j}{RTS}$ $\dot{x}$

Looking at the average TC in stochastic frontier production, the average rate of TC per period was 1.06%. Specifically, the shift of the production frontier function was reduced by 1.55% in 2010, and then increase to 0.28%, 1.99%, 3.50% in 2012, 2014, 2016 respectively. It shows that technical change is quite low in compare with rice produce in Vietnam. In the other hand, Opportunity costs of agricultural production in general and rice production in particular are low. Thus, there is a shift from the agricultural sector to other sectors as well as the internal epidemic in the agricultural sector through shifting cultivation to livestock or some specialty crops with higher value. Technical efficiency change is positively in the period (the average increase of 3.30%) TEC tends to decrease. Specifically, in 2010 increased by 6.75%, in 2012 the growth rate decreased to 3.57%, in 2014 the growth rate decreased to 1.88% and in 2016 the growth rate was only 0.99%.

In the general, there are two problems: (i) The enlargement of TC is expressed through the shift in the direction of stochastic frontier production function; (ii) The positive change of households through the rational use of inputs as well as the combination of inputs to improve technical efficiency. Thereby, narrowing the gap from units to stochastic frontier production function. The combination of both problems leads to an increase in the coefficient of eta and sigma_u2 in the research model.

The estimated TFP growth results are shown at the end of Table 3.9. The average growth of TFP was 4.29%. However, the difference in TFP between years is not too large. When analyzing TFP growth, we see that scale component (SC) plays a small role in TFP growth. Mean while technical change and technical efficiency change play a leading role. Considering the period, the technical change increasingly plays an important role while technical efficiency change is more and more lose its role.

3.4.6. Factors influencing technical efficiency

Wald test was done to ascertain whether the coefficients for the model covariates significantly influenced TE. Based on the p-values, we rejected the null hypotheses that the coefficients for gender, level of education, and ethnicity of the household head, irrigation, disaster, land defragmentation were not equal to zero, meaning that including these variables creat a statistically significant improvement to the model fit.

Table 3.10. Factors influencing on technical efficiency

Variables	Coef	Std. Err.	[95% Conf. Interval]	
Z_1	0.0116193***	0.0033224	0.0051075	0.0181310
Z_2	0.0060172***	0.0012838	0.0035009	0.0085334
Z_3	0.0107987***	0.0015699	0.0077218	0.0138756
Z_4	-0.0137197***	0.0027639	-0.0191369	-0.0083025
Z_5	0.0137918***	0.0006559	0.0125062	0.0150774
Z_6	0.0007164	0.0004704	-0.0001955	0.0016484
Z_7	0.0119039***	0.0036712	0.0047084	0.0190993
Constant	0.8266323	0.0060728	0.8147298	0.8385348

(Source: Vietnam Access to Resources Household Survey)

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

Table 3.10 presents a summary of the drivers of TE. Most of the variables in the model are statistically significant at 99% confidence interval (except Z_6). The coefficient of Z_1 variable is positive indicating that TE of women-headed households is higher than of male-headed households. This reflects the real situation in Vietnam: men usually plays the most important role in the family (household leader). Thus, they are more accessible to public and social resources. Therefore, men often make better decisions than women. In addition, female headed households are more likely to be single or older and do not have much resources for rice production.

The coefficient of Z_2 and Z_3 are positive, indicating that level of education and highest degree of the household head positively influences TE. This is plausible, given that education and highest degree facilitates receptiveness to new ideas and thus adoption of new technologies in rice production thereby increasing TE. This result is similar to some previous research results (Hien et al., 2003; Hoang Linh, 2012; Khai & Yabe, 2011; Le et al., 2017; Tuan M.Cao et al., 2017).

The coefficient of Z_4 is negative, explaining that the Kinh-headed or Hoa-headed households have lower TE than other ethnicities. It sounds counterintuitive because the Kinh and Hoa people often have better access to society. However, it is because of the better access, those household has shifted to other jobs, other crops in the industrialization. It leads to the redundancy for rice. Eventually, in some places, many households of Kinh or Hoa fallow agricultural land. Meanwhile, for other ethnic groups, rice is still the important crop so it is possible that TE of other ethnic households is higher than it is in Kinh or Hoa households.

The coefficient of Z_5 is positive, meaning that irrigation positively influenced TE that the proportion of irrigation takes a positive impact on TE. This is true in reality and confirms the very important role of irrigation in rice production in Vietnam. This result is similar to some previous research results (Hoang Linh, 2012; Khai & Yabe, 2011).

The coefficient of Z_7 is positive. It explains that TE in households with less land fragmentation is lower than TE than those with land fragmentation. This result is contrary to previous research results (Van Hung, MacAulay, & Marsh, 2007). This seems to be contrary to reality. When households have little fragmentation, the application of mechanization into rice production is easier, leading to high efficiency. In addition, less fragmentation also reduces the other costs, decrease labour, cultivation and harvesting activities become easier. However, for agriculture in general and rice production in particular, it depends greatly on natural conditions. Typically, the difference between RRD and MRD. Although there is large fragmentation in RRD, rice yields are much higher than MRD and TE in RRD is higher than MRD. When conducting a small scale study of the same size and productivity, less fragmented households often have higher TEs than households with large land fragmentation. This is in line with the policy of the state in land consolidation and land conversion to reduce land fragmentation, making advantages for high investment to improve TE.

3.5. Conclusions and recommendations

This study used MLE method with balanced panel data to estimate TE. Based on different assumptions, eight models were used in this study. LR test was used to select the best models. The TE mean in the whole Vietnam is 92.62%. Thus, with the same level of inputs and technologies, it is possible to increase TE by 7.38%. TE in rice production improved during the study period. The average TE increased from 82.67% in 2008 to 98.46% by 2016. There are TE differences among the six economic sectors, the highest is the North Central Coast, followed by the Red River Delta and the lowest is the Central Highland. All input variables have positive impact on the quantity of rice production of households. In which, cultivated land is the most important input factor, followed by fertilizer, seed, labour, other costs and pesticide, herbicide. RTS analysis showed that a 1% increase of inputs increased quantity of output by 0.92% during the study period, RTS tends to increase over the years 0.77% (2008) to 0.97% (year 2016). Analysis of technical change shows that the production frontier function of rice production tends to increase 1.06% in each period. From 2006 to 2016, TFP growth is 4.29%. There was no significant difference between each period

Tobit model estimated to predict the drivers of TE with social-characteristics and other specific variables as predictors. The results show that TE is positively influenced by gender, education level, highest degree of the household head, irrigation index land fragmentation proxied by the defragmentation index but negatively influenced by ethnic of household head.

Therefore, in order to increase the TE of rice farmers, the government should address the following issues: (1). There should be a policy of land consolidation to increase the land area under rice; (2). Legislation and incentive mechanisms should be created to facilitate the consolidation of land plots so that farmers are more likely to cultivate on large fields; (3). It is necessary to have appropriate investment policies for irrigation systems in some areas with unfavorable climatic conditions; (4). Strengthening of extension services and vocational training to boost farmer capacities in applying new technologies to rice production is important; (5). Farmers need to be proactive in the prevention of diseases and the impacts of the external environment on rice farming, optimum combinations of inputs for rice production should be selected, particularly cutting down on inputs such as seeds and pesticides and herbicides.

4. Technical inefficiency and production risk of maize farming: A Case study of the Northwestern, Vietnam

Abstract

The determinants of production risk and technical efficiency are able to be studied based on a combination of two methods: Just and Pope's stochastic production and stochastic frontier analysis. Experimental analysis was done by using the panel balance data of 435 maize producing households, which were collected every 2 years from 2008 to 2016 (2008, 2010, 2012, 2014 and 2016), through 5 surveys, 2175 observations. By using the log-likelihood ratio test, the most the optimal model was selected. Estimation marginal output risk provides that land, labour, pesticide, and herbicides raise the possibility of output variances, while seed, fertilizer and other costs reduce the output variances. Determinants of technical inefficiency showed that gender of the household leaders, the number of household members take positive impact on the technical inefficiency. Conversely, education and highest degree of the household leaders, irrigation, disaster, and land fragmentation have a negative impact on technical efficiency. The average TE of maize production in the North West is 82.75%. The average value of TE increased steadily over the period and by 1-2%. Dien Bien is a province that gains the highest average TE, followed by Lao Cai and the lowest is Lai Chau. Research gives some recommendations to increase maize production, eliminate technical inefficiencies, and minimize the impact of risk during production.

4.1. Introduction

4.1.1 Background and problem statement

In order to measure technical efficiency, there has been an assumption that not all the firm are fully technical efficiency. In the literature, technical efficiency measurement is divided into two types: non-parametric method and parametric method. Data envelopment analysis (DEA) is one of the most important methods to measure non-parametric. This method estimates by linear programing and there is no need to use production function or cost function. On the contrary, stochastic frontier analysis (SFA) is on of the methods to measure parametric method. By this method, it is needed to use production function and cost function. The main advantage of SFA is that it deals with random interference. Simultaneously, it is possible to do statistic checks on the structure of production and the level of inefficiency. This is a feature that the DEA method can not do (T. J. Coelli, Rao, O'Donnell, & Battese, 2005).

Traditional SFA studies used to assume that the error term is homoskedastic. Thus, these studies focused on the reasons that impacted on output, estimated technical efficiency and the explained factors that affect on technical inefficiency. Therefore, those studies took little concern about the effect of input factors on output variance. In fact, heteroskedaticity may appear with the error component and affects the estimation of technical efficiency (Kumbhakar & Lovell, 2003). Therefore, when studying the technical efficiency, it is necessary to consider whether the SFA model is consistent with the heteroskedasticity model.

Production risk is an inherent feature of most biological and agriculture production. One of the basic characteristics of production risk is the amount of input affects on output risk. In other words, some inputs make increase output risk while others decrease (Tveteros, 1999). There are compelling reasons for taking this risk into account in empirical analysis of firm behavior and productivity change. First, risk-averse producers choose input levels which differ from the optimal input levels of risk-neutral producers. Second, risk averse producers will be concerned about risk properties when they consider adoption of new technologies, and may not necessarily choose the technology with the highest mean output (Tveteros, 1999).

The determinants of production risk and technical efficiency are able to be studied based on a combination of two methods: Just and Pope's stochastic production and a stochastic frontier analysis (Battese, Rambaldi, & Wan, 1997; Bokusheva & Hockmann, 2006; Jaenicke et al., 2003; Kumbhakar, 2002; Ligeon, Jolly, Bencheva, Delikostadinov, & Puppala, 2013; Ogundari & Akinbogun, 2010; Oppong, Onumah, & Asuming-Brempong, 2016; Tiedemann & Latacz-Lohmann, 2013; Villano & Fleming, 2006).

Maize is an important crop in the economy of some countries and it is grown throughout the world. With the United States, China, and Brazil being the top three maize-producing countries in the world, producing approximately 563 of the 717 million metric tons/year. Maize can be processed into a variety of food and industrial products, including starch, sweeteners, oil, beverages, glue, industrial alcohol, and fuel ethanol (Ranum, Peña-Rosas, & Garcia-Casal, 2014). There are a number of studies using parametric methods to assess the technical efficiency of maize production in different countries around the world (Addai & Owusu, 2014; Ahmed, 2014; Anupama, Singh, & Kumar, 2005; Aye & Mungatana, 2010; Chirwa, 2007; Dang, 2017; Dlamini, Masuku, & Rugambisa, 2012; Essilfie, Asiamah, & Nimoh, 2011; Ghulam, Farman, Khalid, Musawar, & Inamullah, 2009; Isaac, 2011; Kibaara & Kavoi, 2012; Kuwornu, Amoah, & Seini, 2013; Mango, Makate, Hanyani-Mlambo, Siziba, & Lundy, 2015; Oppong et al., 2016; Wakili, 2012). In most cases, these studies used traditional SFA estimation method proposed by Aigner et al. (1997). Thus, it was not possible to explain the risk in the production process. This may lead to bias in TE estimation (Villano & Fleming, 2006). With the traditional SFA estimation, those studies focused only on factors affecting maize yield, TE estimation of maize and found the factors that affect on technical inefficiency (or TE). Some studies estimated output elasticity and return to scale.

In Vietnam, due to economic growth and urbanization, the demand for meat products has increased significantly. Consequently, maize becomes the second most important crop (the first is rice) and the main source of food supply for cattle and poultry. Maize also contributes to reducing poverty in rural areas, lessening deforestation and land degradation (Dinh Thao, Tri Khiem, Xuan Trieu, Gerpacio, & Pingali, 2004; Keil, Saint-Macary, & Zeller, 2009).

The Northern mountain areas is the agro-ecological area that provides most of Vietnam's maize production. Of which, the North East is mostly midland and mountainous with an average elevation of 400-500 meters and the Northwest is the upland and mountainous with elevation from 700-2000 meters. Maize is grown in the northern highlands with an average of 1.5 ha per household, which plays an important role in the household economy.

So far, there is no research on the technical efficiency of maize production in the northern mountainous areas of Vietnam, where the largest production in Vietnam has been made. In particular, there has been no serious examination of the risks of production related to the utilization of inputs in maize production in particular and in the agricultural sector in general. Evaluation of the effectiveness of maize farming households in combination with production risk will make TE measurement more accurate, by evaluating the effect of input use on output variance (production risk), it will elicit important policy implications in agricultural development planning (Jaenicke et al., 2003; Villano & Fleming, 2006). Therefore, the study of technological efficiency and production risk in maize production in northern Vietnam has become urgent to address the following questions: (i) What are the production risks of inputs utilization?; (ii) How is the technical efficiency of maize production in the northern mountainous of Vietnam?; (iii) What factors affect on technical inefficiency of maize production in northern Vietnam?

In this study, the technical inefficiency analysis was extended by taking into account the production risk of maize farming households in the northern mountains. Those households have to face difficulties from economic, cultural, disasters, irrigation, household size, land fragmentation, etc. These risks play important role in the decision making process of households to allocate and utilise inputs. It means they also impact on their output.

4.1.2. Objectives

The main objective of this paper is to analyze the production risk and technical efficiency of maize production in the northern mountainous region of Vietnam. In details:

- To select the best model for estimating maize yield with different inputs;
- To estimate production risk by inputs;
- To estimate technical efficiency of maize farming households in the Northern Mountains;
- To identify factors that affect on technical inefficiency.

4.1.3. Justification of the study

- Justification for input with mean output: Estimation of inputs to output will provide an overview of the relationship between inputs and outputs. The estimation of the output elasticity with the input will show how the output changes when an input changes and the other elements are constant. The scale elasticity estimate will show the change in output if all inputs change at the same rate. This estimate suggests policy related to appropriate inputs leading to increased maize yield.

- Justification for input bias with output variance: When the risky component appears in the structure of SFA function, analysis this case gives a detailed view of how the individual inputs impact on output variance. As a result, there may be some input factors increase output variance and some factors reduce it. This information recommends maize farming households with a more reasonable allocation of inputs.

- Justification for TI: The level of technical inefficiency provided indicates the level of technical utilization and potential of improvement. Factors affecting TI are identified as good bases for recommendations to the government to improve policies related to maize production and to advise maize farming households on improving TE of maize production in the Northern of Vietnam.

4.2. Literature review

4.2.1. Theoretical background of Stochastic frontier and the combination of production risk in the stochastic frontier model

SFA takes its source from Meeusen and van den Broeck (1977) and Aigner, Lovell, and Schmidt (1977). Both articles were published adjacently and there are a few similarities. Shortly thereafter appeared a third paper written by Battese and Corra (1977). Model is expressed as:

$$Y = f(x;\beta).\exp(V-U) \tag{1}$$

where Y is scalar output, x is a vector of inputs, β is a vector of technology parameters, the symmetric random error V accounts for statistical noise and production risk. All three articles focus on analyzing structural error. The first error in the structure is $V \sim N(0, \sigma_v^2)$ to represent the effect of statistical noise. The second error in the structure is $U_i \geq 0$ for the effects of technical efficiency. The difference between them is the believing on the distribution of U_i, the distribution of U conformed to an exponential distribution (Meeusen & van den Broeck, 1977). It was a half-normal distribution (Battese & Corra, 1977). Meanwhile, it could have both types (Aigner et al., 1977). A few years later, Green (1980a, b) discovered Gamma distribution and Stevenson (1980) found out Gamma and truncated normal distribution. (Schmidt & Lovell, 1979) completed the Cobb-Douglas function.

Most studies use output-oriented to measure technical efficiency by the ratio observed output to the corresponding stochastic frontier output. The stochastic frontier output varies from the deterministic frontier by noise affect. And technical efficiency value is between 0 and 1.

$$TE = \frac{Yi}{\exp(x_i'\beta + Vi)} = \frac{\exp(x_i\beta + Vi - Ui)}{\exp(x_i'\beta + Vi)} = \exp(-U_i) \quad (2)$$

Cross-sectional data only provides the most common on producers and the efficiency of producers. The panel data gives us a more general and detailed picture of their transition over time periods. In the version panel data, it is able to be formed generally as follows:

$$Y_{it} = f(X_{it};\beta)e^{Vit-Uit} \qquad Uit > 0;\ i = 1...N;\ t = 1...T \quad (3)$$

There are 2 types of relationship between technical inefficiency and time, they are time-variant and time-varying model. Time-variant model assumes that the technical inefficiency is fixed over time (U_i) (Battese & Coelli, 1988; Schmidt & Sickles, 1984) while time-varying model allows technical inefficiency to change over time (U_{it}) (Battese & Coelli, 1992; Cornwell et al., 1990; Kumbhakar, 1990).

The shortcoming of stochastic frontier production function is its ability to describe the appropriate technology. Just and Pope (1978) pointed out that the effect of inputs on output is not tied to the effects of inputs on output variance. Therefor, if an input has a positive effect on the output and then a positive effect of the input to the variation of the output needs to be applied. To solve that problem, a stochastic specification was introduced in order to generalize traditional production functions. At this point, the production function consists of two component functions: one is the function that specifies the effect of the input to the output mean and the other is to indicate the effect of the input on the output variance. Battese et al. (1997) identified a stochastic frontier production function with the addition of heteroskedastic in the error structure. So we have:

$$Y_{it}=f(X_{it};\beta) + g(X_{it};\psi)(V_{it} - U_{it}) \quad (4)$$

In which: Y_{it} is the output which was produced out of the i-*th* farm in t-*th* year, $f(X_{it};\beta)$ is the mean output function, $g(X_{it};\psi)$ is the output risk function, X_{it} is the vector of inputs, β is the vector of parameters to be estimated in the mean production function, ψ is the vector of parameters to be estimated in the variance production function , V_{it} is random errors and is assumed as independent and identically distributed $N(0, \sigma^2_v)$, U_{it} is assumed as independent and identically distributed non-negative truncations of the $N(\mu,\sigma^2)$ distribution or half-normal distribution $N(0,\sigma^2)$

The about function is the characteristics of stochastic frontier production function. The flexible risk was proposed by Battese et al. (1997). The mean output of the i-*th* farm in t-*th* year is:

$$E(Y_{it}|X_{it}, U_{it}) = f(X_{it}; \beta) - g(X_{it}; \psi)Uit \tag{5}$$

The variance of output or production risk is defined as follows:

$$Var(Y_{it}|X_{it}, U_{it}) = g^2(X_{it}; \psi) \tag{6}$$

The marginal production risk with respect to the j-*th* input is defined to be the partial derivative of the variance of production with respect to X_j

$$\frac{\partial Var(Y_{it}|X_{it}, U_{it})}{\partial X_{ijt}} = \frac{\partial g^2(X_{it}; \psi)}{\partial X_{ijt}} = 2g_i(X_{it}; \psi)g_{ij}(X_{it}; \psi) \tag{7}$$

Thus, $\frac{\partial g^2(X_{it};\psi)}{\partial X_{ijt}} > 0$ => Risk increasing of the j'th input and vice versa. Hence, marginal effect of the j-th input on production risk can be positive or negative depending on the sign of $g_i(X_{it}; \psi)$ and $g_{ij}(X_{it}; \psi)$. In which, $g_{ij}(X_{it}; \psi)$ is the partial derivative of the production risk function with the j-th input.

The technical efficiency of the i'*th* farm is given by equation

$$TE_{it} = \frac{E(Y_{it}|X_{it}, U_{it})}{E(Y_{it}|X_{it}, U_{it} = 0)} = \frac{f(X_{it}; \beta) - g(X_{it}; \psi)Uit}{f(X_{it}; \beta)} = 1 - \frac{g(X_{it}; \psi)Uit}{f(X_{it}; \beta)} = 1 - TI_{it} \tag{8}$$

In the multiplicative operation, the technical inefficiency is attached to the mean production function (Kumbhakar, 2002)

$$Y_{it}=f(X_{it};\beta)\exp(-U_{it}) + g(X_{it};\psi)\, V_{it} \tag{9}$$

In the more flexible form suggested by Kumbhakar (2002), where an additional function *q*(z) for explaining technical inefficiency is introduced:

$$Y_{it}=f(X_{it};\beta) + g(X_{it};\psi)\, V_{it} - q(z_{it};\, \delta)U_{it} \tag{10}$$

4.2.2. Some studies on production risk and technical efficiency

Asche (1999) refers to the production risk model with two steps. Firstly, the occurrence of production risk was tested. If production risk occurs, the mean production function and risk function will be separately estimated. This allows flexibility in individual functions. This method was applied to the empirical analysis of Norwegian salmon. This study uses a linear quadratic mean production function in experiment. In the modeling, the mean function indicated that food is the most important factor with the elasticity of the output, followed by fish input, and the returns to scale was 0.89. Research shows that risk composition plays an important role in the production decisions of risk producers, including optimal input levels and the application of new technologies. This is the premise of the two-step approach in research and have been used by many experimental studies.

Bokusheva, R. & Hockmann, H. (2006) conducted research on production risk and technical inefficiency in Russian Agriculture. The study used panel data from 1995 to 2001 with 447 large agricultural enterprises in three regions. The study used Just and Pope model (1978) to estimate a production function considering production risk, and its extension, by incorporating technical inefficiency as specified by Kumbhakar (2002) in the framework of cross-sectional data. Research showed that risk was the main cause of changes in agricultural production in Russia, and indicates that farm growth was not caused by TE increase. Estimates showed that there was a significant change in production and parameters in the production risk function of high value such as labour and capital.

Villano (2006) conducted a study on technical inefficiency and production risk in rice farming: evidence from Central Luzon Philippines. The study used two types of functions, the translog form and the quadratic form applied to panel data with a sample of 46 households, collected in 8 years (1990-1997). Both frontier models showed that the output elasticity reached highest level with the rice area, it was double in comparison with output elasticity of the fertilizer and labour. Research also indicated that area, fertilizer, and labour increased risk, while herbicides are risk averse. Technical efficiencies of both models were not significantly different, highest in 1992 and lowest in 1996.

Ogundari (2010) conducted a study on Modeling Technical Efficiency with Production Risk: A Study of Fish Farms in Nigeria. The sample was 64 fish farms in Nigeria. The study used a quadratic functional form which is flexible in the sense of a second-order approximation of any unknown mean-output, and then CD function was used to define the variance function. Through the estimation of elasticity, research indicated that food contributed the most to production, followed by labour and fertilizers. Simultaneously, research pointed out that when all factors increased 1%, fish yield increased by 0.98%. The study also found that fertilizers and food raised output variance while labour was on the contrary. Regarding technical efficiency, education and market opportunities spread TE of fish farms in the study area.

Tiedemann & Latacz-Lohmann (2013) accomplished research on production risk and technical efficiency in organic and conventional agriculture - the case of arable farm in Germany. Empirical analysis was conducted by using panel data from 1999/2000 to 2006/2007 on 37 organic and normal farms. Analytical results showed that output change in both production technologies was a major cause of manufacturing risk. Land and labour were two inputs that increase the risk, while seed costs and soil quality reduce the risk.

Oppong (2016) undertook a study on the technical efficiency and production risk of maize production in Ghana. The stochastic frontier model with versatile risk characteristics was used for 232 households in the Brong-Ahafo region. The research demonstrated that the translog function was consistent with the mean output function. The results also showed that seed and labour reduced production risk while land and intermediate input costs increased the risk. The study estimated that technical efficiency of 62% and 38% of potential output were lost due to technical inefficiencies and production risk factors.

4. 3. Data, conceptual framework and empirical estimation

4.3.1. Data

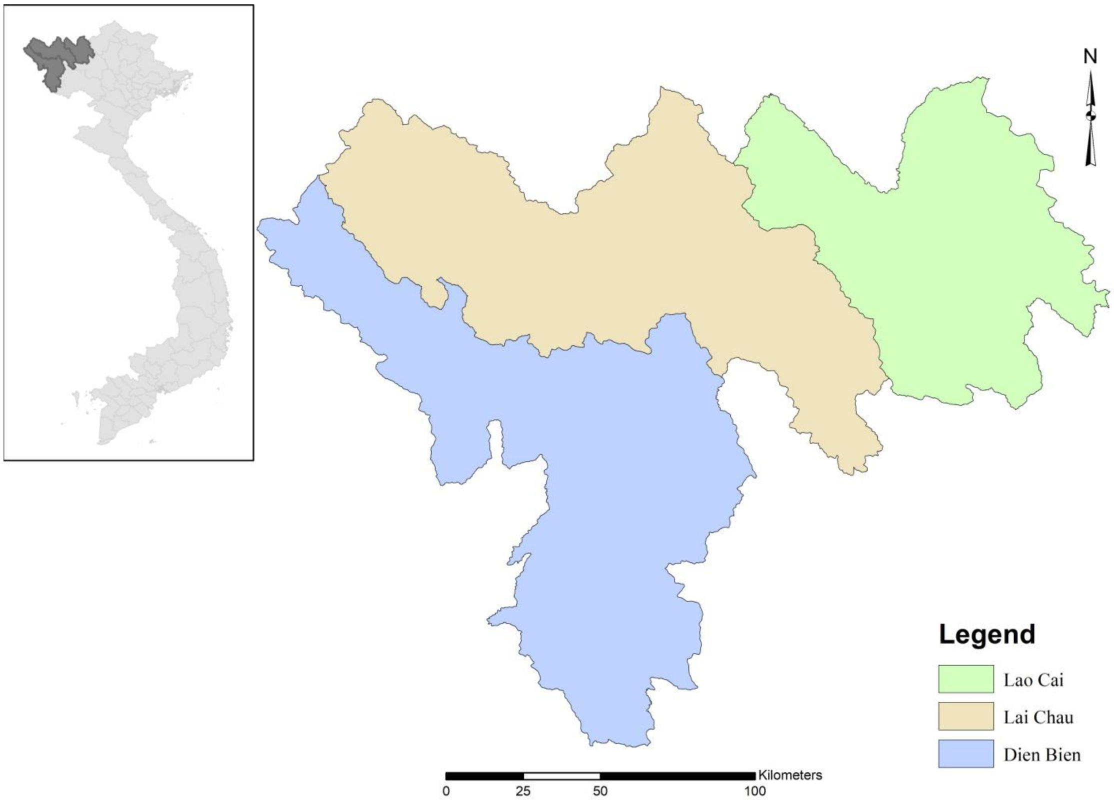


Figure 4.1. Map of research areas

Currently, there are two main data surveys in Vietnam on national scale, those are: 1. Vietnam household living standards survey (VHLSS) since 1992, implemented by the General Statistics Office with support form the World Bank and 2. the Vietnam access to resources household survey (VARHS), implemented by Institute of Labour Science and Social Affairs (Ministry of Labour Invalids and Social Affairs, Viet Nam), funded by Danish International Development Agency (DANIDA). While VARHS has observed the same farmer households in period of time, VHLSS is a census and targets on all households in urban and rural areas.

This paper uses the data of the Vietnam access to resources household survey, which was provided by the Institute of Labour Science and Social Affairs, under the Ministry of Labour Invalid and Social Affairs of Vietnam. The VARHS survey began in 2002 with a small sample (932 households in 4 provinces). In 2006, the sample was increased to 2324 households in 12 provinces. Round 2008 was 3223 households, round 2010 was 3202 households (in which, 2200 panel households). Round 2012 was 3700 households, round 2014 was 3648 households and 3582 households in round 2016. After filtering and connecting data, the panel balance data used in this study including 435 maize producing households, which were collected every 2 years from 2008 to 2016 (2008, 2010, 2012, 2014 and 2016), through 5 surveys, 2175 observations are as follows:

Table 4.1: Observations by regions

Detail	Lao Cai province	Lai Chau province	Dien Bien province
Number of households	196	133	106
Observation	980	665	530

(Source: Vietnam Access to Resources Household Survey)

4.3.2. Conceptual framework

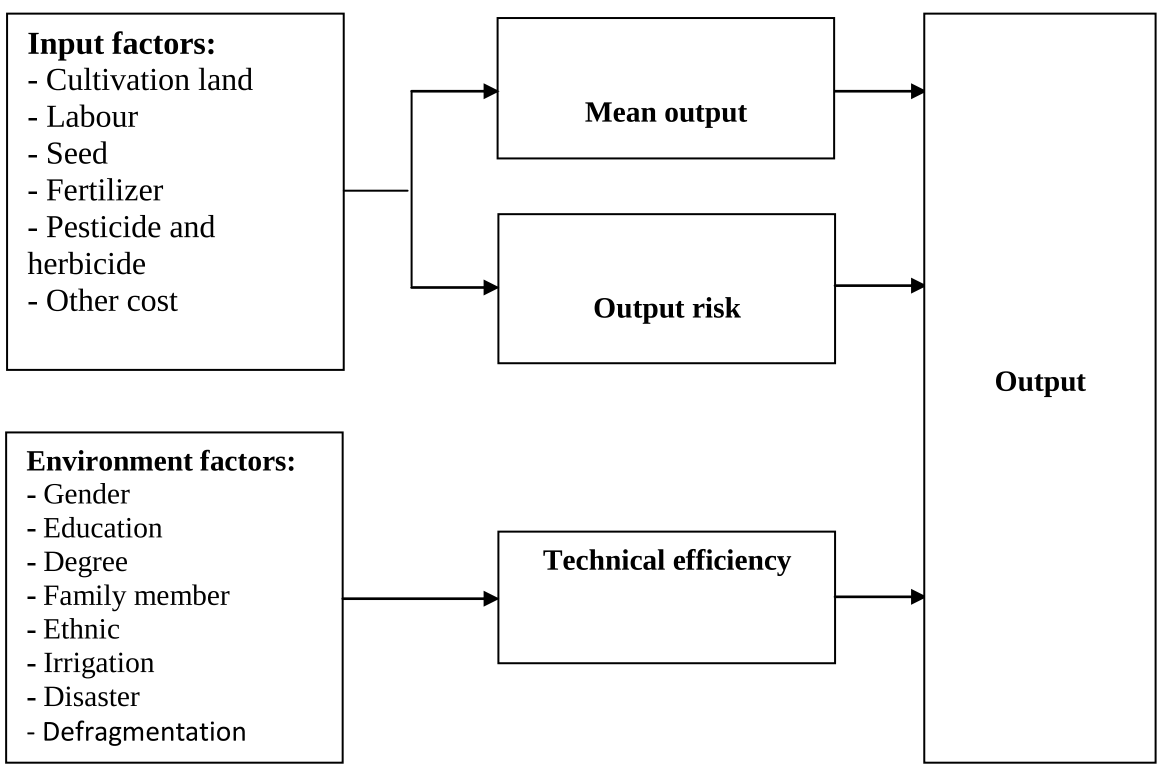


Figure 4.2. Conceptual framework

4.3.3. Empirical estimation

In this study, a flexible form suggested by Kumbhakar (2002) was used, where an additional function *q*(z) for explaining technical inefficiency is introduced:

$Y_{it}=f(X_{it};\beta) + g(X_{it};\psi)\ V_{it} - q(z_{it};\ \delta)U_{it}$

* The mean stochastic frontier production function

In empirical research, it is important to select a structural form of the function as it has a large impact on the estimation of parameters. In the random boundary model, there are usually two forms used: the Cobb-Douglas function (CD) and the translog form. This study was conducted with both forms then using the likelihood ratio test to find the appropriate form function. In addition, the likelihood ratio test was used to test for different assumptions, such as the apportionment of truncated distribution N (μ, σ_u^2) with half-normal distribution N (0, σ_u^2); between time-varying and time-invariant; between the appearance with the absence of technical inefficiency (Uit) from which to choose the best model.

The Cobb-Douglas form of the stochastic frontier production carried out in this study is:

$$\ln Y_{it} = \beta_0 + \sum_{j=1}^{6} \beta_j \ln X_{jit} - U_{it} \quad (11)$$

The translog function form was used as flexible functional and usually use in production reasearches. The translog form of the stochastic frontier production was defined as follows:

$$\ln Y_{it} = \beta_0 + \sum_{j=1}^{6} \beta_j \ln X_{jit} + 0.5 \sum_{j}^{6} \sum_{k}^{6} \beta_{jk} \ln X_{jit} \ln X_{kit} - U_{it} \quad (12)$$

Where the subscripts j, k refer to inputs; i = 1,...,N to firm and t = 1, ...T to years. lnY denotes the natural logarithm of quantity of output, $\ln X_1$ denotes the natural logarithm of area of maize cultivation land, $\ln X_2$ denotes the natural logarithm of total number of days spending on maize production, $\ln X_3$ denotes the natural logarithm of seed/seedling expenditure, $\ln X_4$ denotes the natural logarithm of fertilizer expenditure (chemical and organic fertilizers), $\ln X_5$ denotes the natural logarithm of pesticides and herbicides expenditure, $\ln X_6$ denotes the natural logarithm of all other expenses (hiring labour or livestock, renting agricultural machines, etc).

The output elasticity for input j is defined as follows:

$$E_j^{TL} = \frac{\partial Y_{it}}{\partial X_{jit}} \frac{X_{jit}}{Y_{it}} = \frac{\partial \ln Y_{it}}{\partial \ln X_{jit}} = \beta_j + 2\beta_{jj} X_{jit} + \sum_{j \neq k}^{5} \beta_{jk} \ln X_{kit} \tag{13}$$

Returns to scale (RTS) is equal to the sum of the *j* output elasticity

$$RTS\ (X_{it}) = \sum_{j=1}^{6} E_j^{TL} \tag{14}$$

* Production risk

The previous studies used the CD form function (Bokusheva & Hockmann, 2006; Jaenicke et al., 2003; Kumbhakar, 2002; Villano & Fleming, 2006) to determine the function of the variance. Maximum likelihood estimation was used to determine the parameters in the production risk function with the CD form function.

$$g(X_{it}; \psi) V_{it} = \psi_0 + \sum_{m=1}^{8} \psi_m \ln X_{mit} + \varepsilon_{it} \tag{15}$$

* Technical efficiency effects

After estimating TI, the tobit model will be used to determine the factors that affect on TI with the dependent variable is the TI estimated at the top for both structures

The tobit function is given by:

$TI = \delta_0 + \delta_1 gender + \delta_2 edu\_hl + \delta_3 degree + \delta_4 f\_mem + \delta_5 ethnic + \delta_6 p\_irri + \delta_7 p\_dis + \delta_8 d\_index$

In which, dependent variable: TI is TI that was estimated in step 1; Independent variables are: (i) gender represents gender of household leader (1: male; 2: female); (ii) edu_hl represents education of household leader (1: Never went to school; 2: Primary; 3: Secondary; 4: High education); (iii) degree represents highest degree of household leader (1: No diploma; 2: short-term vocational training; 3: long-term vocational training; 4: professional high school; 5: junior college diploma; 6: Bachelor degree; 7: master degree; 8: PhD); (iv) f_mem represents number of family member; (v) ethnic represents household leader ethnic (1: if Kinh or Hoa; 0 otherwise); (vi) cul_land represent total of rice cultivate land (m2); (vii) p-irri represents the result of dividing the total of irrigated paddy land by total of paddy land (1: irritation <=0.2; 2: 0.2< irrigation<=0.4; 3: 0.4< irrigation<=0.6; 4: 0.8< irrigation<=0.8;5: 0.8< irrigation<=1); (vii) p_dis represents the result of dividing the total of paddy land which was effected by disaster in the observation year by total of paddy land (1: disaster <=0.2; 2: 0.2< disaster <=0.4; 3: 0.4< disaster <=0.6; 4: 0.8< disaster <=0.8; 5: 0.8< disaster <=1); d_index represents defragmentation index. d_index = 1 - sum (a_i^2/A^2). In which, a_i is the size of each plot, A is total farm area. $0 \leq$ d_index ≤ 1. If d_index is approaching to 0, the land defragmentation is not much and vise versa.

4.4. Results and discussion

4.4.1. Descriptive statistics

The means of the independent and dependent variables which was used in the stochastic frontier function are shown in the table below.

The average of maize yield in the northern mountainous region increased from 1123kg (2008) to 1626kg (2016). However, it was unevenly increasing between years and provinces. The average area for maize cultivation was around 3900 square meters per household over the years 2008 to 2014, except 2016, the area for maize cultivation soared to 5354 square meters, of which the Lai Chau increased more than 2 times. The average number of laborers tends to decrease over the years. Meanwhile, the costs of seeds, fertilizers, herbicides and herbicides, and other costs tend to increase over the years.

Table 4.2. Description statistics of some compomemts in maize cultivation

Region	Variable	2008		2010		2012		2014		2016	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
All	Total quantity of output	1,123	1,231	1,108	1,438	1,360	1,253	1,309	1,305	1,626	1,342
	∑ of maize cultivated land	3,990	4,198	3,907	3,854	3,920	3,398	3,934	3,009	5,354	4,377
	Total of labour	105.90	76.79	76.31	65.27	66.98	53.13	50.32	35.49	58.81	42.40
	Seed	97.68	393.10	353.40	380.10	339.60	376.00	750.07	874.20	858.30	822.70
	Fertilizer	72.23	772.50	863.00	2,285	745.40	968.00	985.80	1,682	1,077	1,211
	Pesticides, herbicides	19.69	98.22	48.58	91.79	46.51	81.26	104.0	164.1	114.0	156.5
	Other costs	113.50	516.70	261.90	668.50	747.70	885.50	1,025	1,115	1,134	1,243
Lao Cai	Total quantity of output	1,175	992.7	1,421	1,873	1,758	1,226	1,414	1,102	1,437	1,072
	∑ of rice cultivated land	3,998	4,448	4,212	4,333	4,508	3,464	3,958	2,647	4,016	2,948
	Total of labour	155.20	82.71	98.24	78.32	81.86	48.94	57.35	40.35	45.40	28.83
	Seed	159.20	571.7	500.20	457.50	511.60	453.20	956.90	1,005	815.30	618.1
	Fertilizer	160.3	1,146	1,733	3,083	1,426	958.5	1,558	2,170	1,325	1,083
	Pesticides, herbicides	1.633	10.78	79.08	111.6	40.77	66.26	70.45	120.70	140.50	198.6
	Other costs	227.8	751.1	403.3	952.9	1,311	995.1	1,432	1,273	1,313	1,223

Region	Variable	2008		2010		2012		2014		2016	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Lai Chau	Total quantity of output	717.4	574.0	609.1	404.5	659.0	630.2	914.0	580.3	1,924	1,485
	∑ of rice cultivated land	3,745	3,788	3,168	2,417	2,687	2,023	3,660	2,554	8,110	5,153
	Total of labour	74.14	41.04	47.40	23.66	37.44	31.08	45.64	26.39	86.86	48.28
	Seed	57.92	75.69	197.30	170.00	148.90	95.97	566.10	465.70	1,147	1,143
	Fertilizer	0	0	32.52	86.55	53.91	116.10	363.00	453.10	1,033	1,500
	Pesticides, herbicides	4.015	14.08	6.955	20.50	20.68	41.76	98.20	93.79	91.97	88.21
	Other costs	30.23	82.54	141.7	178.1	227.5	301.3	866.7	699.8	1,298	1,411
Dien Bien	Total quantity of output	1,535	1,905	1,157	1,151	1,503	1,505	1,610	2,013	1,602	1,535
	∑ of rice cultivated land	4,281	4,230	4,272	4,262	4,380	4,183	4,233	4,004	4,370	3,965
	Total of labour	54.62	32.91	72.04	59.48	76.51	66.31	43.22	33.71	48.41	38.93
	Seed	33.83	106.60	277.80	308.70	261.00	289.50	601.20	930.90	575.20	507.10
	Fertilizer	0	0	297.00	1,145	354.70	724.90	709.60	1,229	675.00	878.70
	Pesticides, herbicides	72.74	188.8	44.42	84.89	89.57	119.0	173.40	257.30	92.69	126.0
	Other costs	6.604	52.13	151.4	229.6	359.0	447.60	471.90	927.50	597.9	846.6

(Source: Vietnam Access to Resources Household Survey)

Table 4.3. Description statistic of some components in tobit model

Variable	*Specific*	*2008*		*2010*		*2012*		*2014*		*2016*	
		Count	Col %	Count	Col %	Count	Col %	Count	Col %	Count	Col %
gender	Male	420	96.55	418	96.09	414	95.17	404	92.87	405	93.10
	Female	15	3.45	17	3.91	21	4.83	31	7.13	30	6.90
Edu_hl	Does not attend school	213	48.97	217	49.89	230	52.87	238	54.71	192	44.14
	Primary school	146	33.56	143	32.87	127	29.20	111	25.52	122	28.05
	Secondary school	59	13.56	66	15.17	66	15.17	71	16.32	95	21.84
	High school	17	3.91	9	2.07	12	2.76	15	3.45	26	5.98
degree	No diploma	429	98.62	420	96.55	405	93.10	402	92.41	413	94.94
	Short-term vocational training	1	0.23	9	2.07	20	4.60	22	5.06	14	3.22
	Long-term vocational training	1	0.23	1	0.23	2	0.46	4	0.92	2	0.446
	Professional high school	4	0.92	5	0.15	8	1.84	7	1.61	6	1.38
	Junior college diploma	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
	Bachelor degree	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
f-mem	Number of family member	6.14		5.93		5.93		5.83		5.65	
Ethnic	Ethnic minority	424	97.47	426	97.93	427	98.16	426	97.93	425	97.70
	Kinh and Hoa	11	2.53	9	2.07	8	1.84	9	2.07	10	2.30

Variable	*Specific*	*2008*		*2010*		*2012*		*2014*		*2016*	
		Count	Col %	Count	Col %	Count	Col %	Count	Col %	Count	Col %
P_irri	<=20%	408	93.79	390	89.66	360	82.76	353	81.15	303	69.66
	20-40%	2	0.46	7	1.61	9	2.07	8	1.84	9	2.07
	40-60%	4	0.92	4	0.92	15	3.45	5	1.15	8	1.84
	60-80%	1	0.23	4	0.92	4	0.92	5	1.15	7	1.61
	>80%	20	4.60	317	6.90	47	10.80	64	14.71	108	24.83
P_dis	<=20%	272	62.53	106	24.37	120	27.59	130	29.89	136	31.26
	20-40%	3	0.69	6	1.38	7	1.61	5	1.15	7	1.61
	40-60%	4	0.92	3	0.69	8	1.84	5	1.15	13	2.99
	60-80%	5	1.15	3	0.69	12	2.76	5	1.15	7	1.61
	>80%	151	34.71	317	72.87	288	66.21	290	66.67	272	62.53
d_index	Defragmentation index	0.23		0.21		0.22		0.25		0.34	

(Source: Vietnam Access to Resources Household Survey)

The mean of the factors affecting TI used in the Tobit function is shown in table 4.3.

Data in the table indicates quite clearly characteristics of the mountainous households in the north of Viet Nam. More than 93% of household leaders are male. This percentage has decreased but not significantly over the years. The education level of household leaders is quite low, most of them have not attended to school (accounting for over 44%). Number of people who completed high school is very low (less than 6%). About the degree of household leaders, most of them are not graduated from high school (accounting for 95%), the rate of participation in short-term as well as long-term training courses is negligible. The number of family members is higher than other regions in the country, this average is 5.6 persons per household. Most of them are in ethnic minority groups (accounting for over 97%) while the Kinh or Hoa account for only a small number (less than 3%). Irrigation rate in upland areas is very low, in 70% of interviewed household, this number is less than 20%. Whereas, disaster rate is quite high, it was more than 80% in 60% of households. The rate of land fragmentation has varied over the years but is not significant.

4.4.2 Production frontier estimates

Maximum-likelihood estimates were used to estimate parameters. 5 models were proposed based on different assumptions. The variables included in the model are statistically significant at over 90% (except for other costs). For CD form (model 3), the inputs have a positive effect on output. For the translog form, it was shown that the cultivated area and fertilizer took negative impact whereas the seed, labour, herbicide, and insecticide did positive effect on yield.

Table 4.4. Maximum likelihood Estimates of stochastic frontier model

Variables	Model 1 (Battese and Coelli 1992)	Model 2 (Mu=0)	Model 3 (B_{ij}=0)	Model 4 (Mu= Gamma =0)	Model 5 (Mu= Eta =0)
lncul_land	-0.390**	-0.375**	0.652***	-0.394**	-0.330**
	(0.154)	(0.154)	(0.0167)	(0.154)	(0.153)
lnlabour	0.683***	0.653***	0.194***	0.611***	0.598***
	(0.126)	(0.125)	(0.0174)	(0.124)	(0.125)
lnseed	0.205***	0.181***	0.0195***	0.167***	0.167***
	(0.0665)	(0.0652)	(0.00731)	(0.0636)	(0.0643)

Variables	Model 1 (Battese and Coelli 1992)	Model 2 (Mu=0)	Model 3 (B_{ij}=0)	Model 4 (Mu= Gamma =0)	Model 5 (Mu= Eta =0)
lnfertilizer	-0.0794**	-0.0906**	0.0293***	-0.0963**	-0.0836**
	(0.0402)	(0.0406)	(0.00431)	(0.0410)	(0.0410)
lnpest_herb	0.0881*	0.0882*	0.0146***	0.0955**	0.0961**
	(0.0455)	(0.0456)	(0.00456)	(0.0455)	(0.0461)
lno_cost	-0.00127	0.0230	-0.00199	0.0442	0.0403
	(0.0599)	(0.0591)	(0.00640)	(0.0575)	(0.0584)
lncul_land_sq	0.171***	0.160***		0.159***	0.152***
	(0.0243)	(0.0240)		(0.0236)	(0.0237)
lnlabour_sq	-0.0405	-0.0476		-0.0486	-0.0427
	(0.0358)	(0.0355)		(0.0348)	(0.0353)
lnseed_sq	0.0661***	0.0625***		0.0657***	0.0683***
	(0.00906)	(0.00896)		(0.00886)	(0.00884)
lnfertilizer_sq	0.0306***	0.0291***		0.0225***	0.0252***
	(0.00462)	(0.00477)		(0.00482)	(0.00484)
lnpest_herb_sq	0.0339***	0.0317***		0.0275***	0.0317***
	(0.00737)	(0.00742)		(0.00741)	(0.00755)
lno_cost_sq	0.0124*	0.0106		0.00878	0.00875
	(0.00677)	(0.00676)		(0.00664)	(0.00671)
lncul_landlnlabour	-0.0393*	-0.0303		-0.0258	-0.0263
	(0.0237)	(0.0234)		(0.0228)	(0.0231)
lncul_landlnseed	-0.0401***	-0.0320***		-0.0285***	-0.0311***
	(0.0102)	(0.00991)		(0.00947)	(0.00961)
lncul_landlnfertilizer	0.00561	0.00716		0.0108	0.00986
	(0.00727)	(0.00730)		(0.00723)	(0.00730)
lncul_landlnpest_herb	-0.00871	-0.00932		-0.0122	-0.0123
	(0.00810)	(0.00810)		(0.00803)	(0.00812)
lncul_landlno_cost	-0.00400	-0.00675		-0.00702	-0.00562
	(0.00941)	(0.00914)		(0.00880)	(0.00895)
lnlabourlnseed	-0.00490	-0.0114		-0.0163	-0.0135
	(0.0114)	(0.0112)		(0.0108)	(0.0110)
lnlabourlnfertilizer	0.00106	0.00260		0.00272	0.00121

Variables	Model 1 (Battese and Coelli 1992)	Model 2 (Mu=0)	Model 3 (B_{ij}=0)	Model 4 (Mu= Gamma =0)	Model 5 (Mu= Eta =0)
	(0.00729)	(0.00726)		(0.00715)	(0.00727)
lnlabourlnpest_herb	-0.00682	-0.00556		-0.000611	-0.000863
	(0.00863)	(0.00865)		(0.00851)	(0.00864)
lnlabourlno_cost	0.00326	0.00364		0.000828	-0.000182
	(0.0101)	(0.00994)		(0.00968)	(0.00980)
lnseedlnfertilizer	-0.0150***	-0.0157***		-0.0165***	-0.0171***
	(0.00418)	(0.00414)		(0.00409)	(0.00412)
lnseedlnpest_herb	-0.00900**	-0.00848**		-0.00800**	-0.00780**
	(0.00391)	(0.00384)		(0.00375)	(0.00381)
lnseedlno_cost	-0.0118***	-0.0102***		-0.00820***	-0.0083***
	(0.00329)	(0.00324)		(0.00309)	(0.00312)
lnfertilizerlnpest_herb	0.00103	0.000998		0.000653	0.000621
	(0.00178)	(0.00180)		(0.00180)	(0.00183)
lnfertilizerlno_cost	0.00676**	0.00682**		0.00697**	0.00678**
	(0.00327)	(0.00326)		(0.00323)	(0.00327)
lnpest_herblno_cost	-0.00398	-0.00359		-0.00234	-0.00405
	(0.00345)	(0.00343)		(0.00339)	(0.00344)
Constant	3.269***	3.385***	0.811***	3.616***	3.238***
	(0.565)	(0.573)	(0.108)	(0.585)	(0.571)
Lnsigma2	-1.784***	-1.741***	-1.510***	-1.232***	-1.528***
	(0.0328)	(0.0549)	(0.0553)	(0.0325)	(0.0504)
Ilgtgamma	-2.631	-1.976***	-1.305***	0	-0.978***
	(0)	(0.426)	(0.261)	(0)	(0.202)
Mu (μ)	-0.198	0	0	0	0
	(0)	(0)	(0)	(0)	(0)
Eta (η)	0.479***	0.235***	0.141***	0.00529	0
	(0.0273)	(0.0610)	(0.0346)	(0.0259)	(0)
$\sigma_S^2 = \sigma_u^2 + \sigma_v^2$	0.1679	0.1754	0.2210	0.2916	0.2170
Gamma (γ) = σ_u^2/σ_S^2	0.0672	0.1218	0.2133	0.5000	0.2732
σ_u^2	0.0113	0.0214	0.0471	0.1458	0.0593
σ_v^2	0.1566	0.1540	0.173	0.1458	0.1577

Variables	Model 1 (Battese and Coelli 1992)	Model 2 (Mu=0)	Model 3 (B_{ij}=0)	Model 4 (Mu= Gamma =0)	Model 5 (Mu= Eta =0)
Log likelihood	-1179.0474	-1177.1541	-1321.4411	-1204.5596	-1187.9131

(Source: Vietnam Access to Resources Household Survey)

Standard errors in parentheses

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

The likelihood ratio test (LR) was used to select the most appropriate model. As usual, LR test should be taken into account to find out if translog production frontier form or the CD production frontier form is the best. However, to minimize the steps, the LR test was used to test between the translog production frontier form and then test for the CD production frontier form.

Take model 2 as the root for comparison with other models with different assumptions. There were some assumptions in model 2 as follows: translog production frontier from, Ui has a half-normal distribution ($U_i \sim N(0, \sigma_u^2)$), assuming time-varying decay and existence of technical inefficiency. (i) In comparison with model 1: LR test results show that the null hypothesis (H_0: $\mu \neq 0$) is rejected, but the result with H_0: $\mu \neq 0$ is mising the p-value of Mu (μ) and inverse of loggamma. Thus model 2 is chosen in order to continue to compare with other models; (ii) In comparison with model 3: The LR test results showed that the null hypothesis (H_0: $\beta_{ij} = 0$) is rejected, it means the model assuming translog production frontier form is more appropriate than the model assuming CD production frontier form, thus model 2 is better than model 3; (iii) In comparison with model 4: The LR test results show that the null hypothesis (H_0: $\mu = \gamma = 0$) is rejected, that is to say the model assuming distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), time-varying decay with the existence of the U_i is more applicable than model with the assumption that distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), time-varying decay and without the existence of the U_i the model (the observed variables are fully technically efficient), thus model 2 is better than model 4; (iv) In comparison with model 5: The LR test results show that the null hypothesis (H_0: $\mu =$

$\eta = 0$) is rejected, that is to say the model assuming distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$), time-varying decay with the existence of the U_i is more applicable than model with the assumption that distribution of U_i is half-normal ($U_i \sim N(0, \sigma_u^2)$) and time-invariant and with the existence of the U_i the model, thus model 2 is better than model 5.

Table 4.5. Hypothesis test for model specification and statistical assumptions

Null Hypothesis	Model	Likelihood Ratio test (LR)	df	Prob > chi2	Decision
H_0: $\mu = 0$	Model 1.2 vs model 1.1	3.79	1	0.0517	Reject H_0
H_0: $\mu = 0$ and $\beta_{ij} = 0$	Model 1.2 vs model 1.3	288.57	21	0.0000	Reject H_0
H_0: $\mu = \gamma = 0$	Model 1.2 vs model 1.4	54.81	1	0.0000	Reject H_0
H_0: $\mu = \eta = 0$	Model 1.2 vs model 1.5	21.52	1	0.0000	Reject H_0
$H_{0:}$ $\psi_1 = .. = \psi_6 = 0$	The Wald test	255.78	6	0.0000	Reject H_0
H_0: $\delta_0 = ... = \delta_9 = 0$	The Wald test	307.59	8	0.0000	Reject H_0

After testing the models with the difference in assumption, model 2 was chosen as the best model and it would be used for the next analysis.

The fifth hypothesis was tested to show that there is the risk of inputs in the production process (Opposition hypothesis was eliminated at 0.001). This explains that the inputs in the model are related to the production risk. In the other words, in the stochastic frontier production model with flexible form, there is the existence of production risk in the production process.

The sixth hypothesis was tested also confirmed that hypothesis of technical inefficiencies did not appear in model was denied at a significance level of 0.001, and in the table "Maximum Likelihood Estimates of stochastic frontier model", chosen model is statistically significant with σ_u^2 is expected to be 0.02 and different from zero.

Eta (η) is positive meaning the technical inefficiency decrease with time, in other words, TE tends to increases by the years. In the model, σ_S^2 is estimated at 0.1754, explaining that variance output is caused by technical inefficiencies and random noise. In order to quantify the importance of production risk with inefficiency for observed variance output, the variance of the error tearm can be used as a based. Kumbhakar & Lovell, 2003 gave an output variance calculation that was used for the model with a half normal distributed inefficiency term as follows:

$$\sigma_S^2 = \sigma_{vi}^2 + \mathrm{VarU} = \sigma_{vi}^2 + \left(\frac{\pi - 2}{\pi}\right)\sigma_{ui}^2$$

It can be shown that variance output, which is explained by production risk, is greater than technical inefficiency. Gamma (γ) is the variance ratio, explaining the total variation in output from the frontier level of output attributed to technical inefficiency. Estimated value of γ is 0.12 which indicates that 12% of total variation in maize yield is due to lack of technical efficiency in the study area.

4.4.3. Output elasticity and scale elasticity

Table 4.6. Estimate output elasticities and scale elasticity

Input	Elasticity	Std.Err	z	P>|z|	[95% Conf. Interval]	
Lncul_land	0.6014	0.0183	32.91	0.0000	0.5656	0.6373
Lnlabour	0.1784	0.0188	9.47	0.0000	0.1414	0.2152
Lnseed	0.0667	0.0120	5055	0.000	0.0432	0.0902
Lnfertilizer	0.488	0.0057	8.52	0.000	0.0376	0.0600
Lnperst_herb	0.0010	0.0053	0.20	0.845	-0.0094	0.0115
Lno_cost	0.0037	0.0077	0.49	0.0626	-0.0112	0.0187
Scale elasticity	0.9000	0.0163	55.19	0.000	0.8681	0.9320

(Source: Vietnam Access to Resources Household Survey)

Estimates of elasticity have shown that land is the most important input to the yield, followed by fertilizer, labour, seed and other costs. This means that expanding maize land, increasing fertilizer investment, adding labour, improving seed and so on will all impact on maize yield. If other inputs are constant, 1% increase in cultivated land, fertilizer utilization, labour, seed or other costs, would lead to result of enlargement 0.601%, 0.488%, 0.178%, 0.067%, 0.004% increase in yields, respectively. Pesticides and herbicides also play positive roles on yields. However, these effects are small and statistically significant. Scaling elasticity can be obtained through the total elasticity of the output. Scale elasticity is obtained through estimation of 0.9. This means that if all the inputs are increased by 1%, the output will increase by 0.9%.

4.4.4. Estimates of marginal output risk

Output variance in the production process is explained by inputs, indicating that it relates to information on production risk management. Some inputs reduce risk while others influence on the other side. Researching this issue gives us a scientific basis for using information to stabilize maize yields of households.

Table 4.7. Maximum likelihood estimates of the linear production risk function

Variable	Parameters	Estimates	Std.Err	P>|z|
Constant	Ψ_0	-6.337046	0.0245533	0.000
cul_land	Ψ_1	0.0000264	4.70e-06	0.000
labour	Ψ_2	0.0025162	0.0002542	0.000
seed	Ψ_3	-0.0000941	0.000028	0.001
fertilizer	Ψ_4	-0.0000667	0.000012	0.000
pest_herb	Ψ_5	0.0004426	0.0001096	0.000
o_cost	Ψ_6	-0.0000297	0.0000176	0.092

(Source: Vietnam Access to Resources Household Survey)

The table below shows that land, labour, pesticide and herbicides raise the possibility of output variances. Although these impacts are statistically significant, their effects are not strong. The results of increased risk by land, pesticides and herbicides coincide with a previous study (Oppong et al., 2016). Meanwhile, the result of risk extension by labour is contrary to previous research on maize production in Ghana (Oppong et al., 2016). Seed, fertilizer and other costs reduce the output variances. This explains that the efficient use of seeds, fertilizers, and other costs can be used to lower output variances. A risk averse producer will therefore reduce land, labour, herbicides and plant protection as these inputs cause productivity fluctuations. Simultaneously, the risk averse producer may boost seedlings, fertilizer as well as other costs to cut down output fluctuations.

4.4.5. Determinants of technical inefficiency

The advantage of the Tobit model is that it allows dependent variables to be constrained between certain values. In this study, the technical inefficiency is between zero and one. Through the Wald test that at least one of the predictors of regression coefficients is not equal to zero. Based on the p-value, null hypothesis can be rejected. In other words, the coefficients of gender, edu_hl, degree, f_mem, hl_ethnic, p_irri, p_dis, d_index are not concurrently zero. Thus, these variables create statistically significant improvement in the fit of the model.

Table 4.8. Determinants of technical inefficiency

Variables	Parameters	Estimates	Standard errors
Constant	δ_0	0.196***	0.01632
gender	δ_1	0.0559***	0.01179
edu_hl	δ_2	-0.00976***	0.00321
degree	δ_3	-0.0153***	0.00573
f_mem	δ_4	0.00702***	0.00132
hl_ethnic	δ_5	-0.00631	0.02717
p_irri	δ_6	-0.0111***	0.00141
p_dis	δ_7	-0.00762***	0.00095
d_index	δ_8	-0.0630***	0.00860

(Source: Vietnam Access to Resources Household Survey)

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

All variables in the model were statistically significant at 99% (except hl_ethnic). The coefficient of gender variable and f_mem positive with TI. TI of male-headed households is higher than that of female-headed households. To put differently, TEs of female-headed households are higher than the TEs of male-headed households. This result is quite similar to some previous studies on maize production (Abdulai, Nkegbe, & Donkoh, 2013; Kuwornu et al., 2013; Mango et al., 2015). This seems to be contrary to the reality in Viet Nam, where the majority of household leaders are male, and men are likely to play the most important role in the family. In addition, men are easy to approach public and social resources than women so that, they usually make better decisions. Therefore, further research is needed to clarify the relationship between the gender of household leader and efficiency in maize production in Vietnam.

The results of the Tobit estimation provides that households with higher number of members have higher TI than those with fewer family members. This result is quite similar to some previous studies on maize production (Kuwornu et al., 2013; Mango et al., 2015; Wakili, 2012). Normally, the number of family members related to the number of available labour in the households. However, in the mountainous region of northwestern Vietnam, family size is 5.6 to 6.14 members, of which the major labour is usually the husband, support from wife, children and elderly seems to be unclear. It leads to TE of household with fewer members is higher than TE of households with more members. In addition, large size households have to face to pressure more strongly than small size households when resources are limited. Large size households are often poor households whose production capacities as well as inputs approach are often ineffective.

The coefficient of education and the highest degree of household ownership are negative. This explains that the higher educational level and degree of household leader is, the lower TI they achieve. It means education and degree higher TE in maize production. This result is quite similar to previous research on maize production (Abdulai, Nkegbe, & Donkoh, 2013; Ahmed, 2014; Aye & Mungatana, 2010; Kibaara & Kavoi, 2012; Kuwornu et al., 2013).

The negation of irrigation shows that when irrigation rates increase, it will reduce TI, simplistically, irrigation increase the TE of the maize producers.

The negative of catastrophe indicates catastrophic rate lower TI. That seems to be against reality. However, in the study areas, disasters such as lack of water, impoverished soil seem not to significantly affect on maize. By way of explanation, maize is suitable for the characteristics of the study areas while some other crops may be unsuitable and conflicting.

The negative of land fragmentation in maize cultivation with TI shows that when fragmentation increases, TI decreases. Otherwise stated, as the fragmentation index step up, the TE rises. This seems to be contrary to logic, but it does fit the reality of maize cultivation in the mountains. Maize cultivation in the study area usually conducted in areas with complex terrain, farmers could choose to plant maize on the land they acquired or to select land that would be suitable for growing maize. Therefore, TIs are high if households plant maize on available and normal plots and it will be low if they chose the new and fat plots (habitual shifting cultivation of ethnic minorities).

4.4.6 Technical efficiency index

The estimated TE value of maize farming households in north-west Vietnam ranges from 41.17% to 97.66%. The average TE of northwestern Vietnam is 82.75%. This indicates a 17.25% loss due to inefficient maize production or inefficiency among households in the sample or a combination of both. This results are similar to some previous studies. From the above results, maize farmers can boost their production by 15.27% [1 (82.75 / 97.66)] by improving technical efficiency. For the lowest technically efficient households, it can increase to 57.84% by technical efficiency improvement.

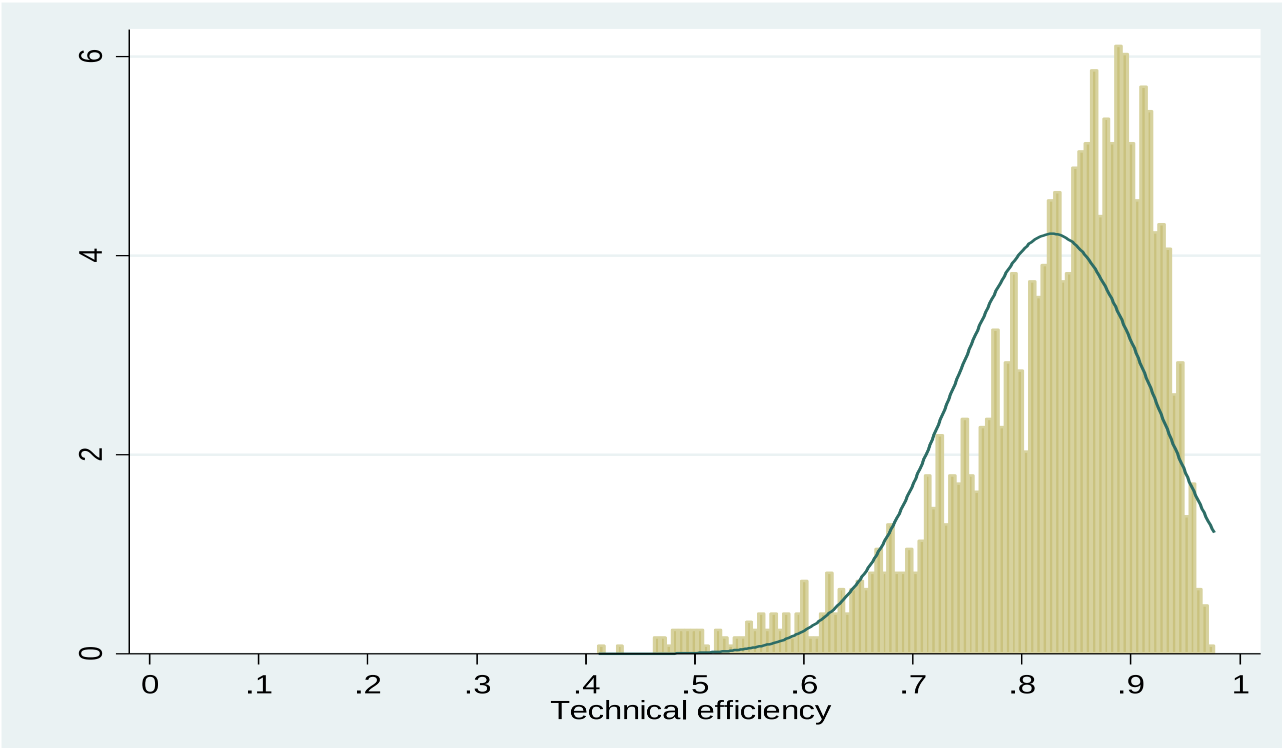


Figure 4.3. Distribution of technical efficiency of maize in Northwest, Vietnam

Technical efficiency is in the range of> = 80% to <90% which is the most concentrated, accounts for 45.66% of total samples (993 households). Followed by TE in> = 90 to = <100%, accounting for 23.26% in the sample size (506 households). TE is in the range of> 70 to 80, accounting for 20.83% of total sample (453 households). TE is in the range of> = 60 to <70, accounting for 6.90% of total sample (150 households). TE ranged from> = 50 to <60, accounting for 2.53% of total sample (55 households). TE ranged from> 40 to <50, accounting for 0.83% of total sample (18 households).

Average TE values by region are shown in the table 4.9. There is a difference in TE between the three provinces, which is explained in the model when Gamma (γ) is estimated to be 0.12. In particular, Dien Bien province has the highest average TE (85.82%), followed by Lao Cai province (84.80%) and Lai Chau province (77.29%).

Table 4.9. Technical efficiency by regions

Regions	N	Mean	Std.Dev	Min	Max
Lao Cai province	980	0.8480	0.0698	0.4303	0.9617
Lai Chau province	665	0.7729	0.1063	0.4666	0.9661
Dien Bien province	530	0.8582	0.0898	0.4117	0.9766
Total	2175	0.8275	0.0946	0.4117	0.9766

(Source: Vietnam Access to Resources Household Survey)

Table 4.10. Technical efficiency by years

Year	N	Mean	Std.Dev	Min	Max
2008	435	0.7509	0.1066	0.4117	0.9413
2010	435	0.7958	0.0910	0.4956	0.9532
2012	435	0.8338	0.0766	0.5740	0.9629
2014	435	0.8655	0.0636	0.6446	0.9705
2016	435	0.8916	0.0523	0.7066	0.9766
Total	2175	0.8275	0.0946	0.4117	0.9766

(Source: Vietnam Access to Resources Household Survey)

The average value of TE over the years is shown in the table 4.10. Data provides that average TE increases over the years with about 1-2%. It was 75.09% in 2008 and 89.16% in 2016. This is also explained in the translog production frontier model with the value Eta (η) = 0.235, inefficiency decreases over time means that TE increases over time.

4.5. Conclusions and recommendations

This paper used a flexible form to study the combination of production risk in the stochastic frontier model. After the test, stochastic frontier model with heteroscedasticity in both is the best fit for the data. Results from an analysis of 435 maize farming households, with 2175 observations during the 2008-2016 period through the MLE method, were used on various assumptions for panel data. At the same time, the stochastic frontier model was chosen as the best that most suitable model. Elasticity of output with inputs is positive. In which, land is the most positive input affect on output, scale elasticity is estimated by 0.9. Variance output which was explained by productive risk is higher than technical inefficiency. MLE is also used to estimate the linear production risk function. Land, labour, pesticides and herbicides have the potential to change output variances, while seeds, fertilizers and other costs have the effect of reducing the variance output. However, these effects are small and negligible. In the Tobit model, gender of the household leaders, the number of household members take positive impact on the technical inefficiency. Conversely, education and highest degree of the household leaders, irrigation, disaster, and land fragmentation have a negative impact on technical efficiency. The average TE of maize production in the North West is 82.75%, of which 45% of the samples are in the range of 80-90%. The average value of TE increased steadily over the years and by 1-2%. Dien Bien is a province that gains the highest average TE, followed by Lao Cai and the lowest is Lai Chau.

Based on the findings of the study, the following recommendations are intended to increase maize production, eliminate technical inefficiencies, and minimize the impact of risk during production.

Maize farming households should invest more in land, labour, seeds, fertilizers, pesticides and herbicides and other costs due to they has the potential to increase maize production. Farmers can transform the crop structure through the conversion of some inefficient paddy fields into maize. These recommendations can be made if the government continues to subsidize and manage inputs such as seed and fertilizer. Expanding fertilizer production in factories, especially NPK. Supporting for import of potassium fertilizer (Vietnam has no potash ore). Providing hybrid maize or genetically modified maize suitable for cultivated land. Continuing to support the development of commodity maize growing areas in this area.

In efforts of reducing technological inefficiencies in maize production, farmers should strive to further improve their education levels while actively seeking short-term training courses in maize production. Farmers should actively prevent the impact of disease and the external environment on maize production, and optimum inputs for maize production under limited resources. The government should have appropriate investment policies for irrigation systems in areas with unfavorable natural conditions, appropriate land policies and good legal corridors to facilitate the consolidation of plots, avoiding land fragmentation. It is necessary to expand the training courses and apply new technologies to production through various channels such as the Women's Union, the Farmers' Association, the non-governmental organizations and so on.

5. General Conclusions

The objective of this dissertation was achieved as presented in 3 papers. The thesis contributed to literature in many different ways through different approaches and objects.

Firstly, through using the maximum likelihood method to estimate the parameter of stochastic frontier of output distance function, the study explored the technical efficiency, technical change, return to scale and TFP of the agricultural sector in Vietnam. Tobit model was used to determine the determinants of technical efficiency of the agricultural sector in Vietnam.

Secondly, providing a suitable model to predict the technical efficiency of rice production in Vietnam. Simultaneously, providing sufficient evidence of the technical efficiency level between the study areas as a good base for policy makers to plan rice production project. Besides, clarifying the factors affecting the technical efficiency of rice production in Vietnam.

Thirdly, through analysis of technical inefficiency and production risk of maize production in Northwest Vietnam, the study explored the impact of the input factors to output risk. Concurrently, proposing solutions to minimize technical inefficiency as well as output risk.

5.1 Major findings

The introductory chapter of this thesis provides the general context of the pressure of population growth, climate change on economies. It also provides an overview of Vietnam's economic and institutional changes that influences Vietnam’s economic development and growth prospects and of the agriculture sector in particular. On this basis, this chapter provides: (1) Vietnam's agricultural outlook as well as the research gap on technical efficiency for this sector in Vietnam; (2) the current situation of rice production in Vietnam, appropriate model selection to analyze variations in TE levels across regions for planning purposes; (3) the current situation of maize production in North West Vietnam, and assessing the impact of level of inputs on output risks. Especially, this chapter emphasizes the importance of improving TE (technical inefficiency) of the farmers.

Chapter 2 focuses on TE, technical change, and return to scale in Vietnam Agriculture in the period 2008-2016, and maximum likelihood estimation is employed on a balanced panel data to estimate parameter of stochastic frontier of output distance function, this study shows that the level of TE in Vietnam agriculture was 89.29%, of which, the highest was observed in Lao Cai province, followed by Lai Chau province and the lastly Phu Tho province. The average technical change for the whole study period decreased by 4.43%. All the inputs positively influenced the value of agricultural output in Vietnam, in which land played the biggest role, followed by intermediate cost and labour. TFP growth grew in almost every period, except in 2012, with an average annual rate of 2.71%. Identifying the factors affecting the technical inefficiency shows that household size and number of plot of the household positively and significantly influenced TE while ethnicity of household head negatively influenced TE.

Chapter 3 zooms in on the level of technical inefficiency in the whole Vietnam and discusses the drivers of TE. The likelihood ratio test aided in selecting the best model. The findings show that TE score of Vietnam was 92.62% and increased over the years. There are TE differences among the six economic regions, the highest is the North Central Coast, followed by the Red River Delta and the lowest is the Central Highland. This research indicates that all inputs have positive impact on the value of rice production in which the most with positive and strongest impact was cultivated land, followed by fertilizer, seed, labour, other costs and pesticide, herbicide. In details, if all other variables are constant, when increases 1% of land, labours, seeding, fertilizer, pesticide and herbicide and other costs, yield value increase 0.57%; 0.08%; 0.09%; 0.1%; 0.03%; 0.04% respectively. Return to scale was 0.93 and increased during the study period. The average rate of TC per period was 1.06%. Technical efficiency change is positively in the period (the average increase of 3.30%) TEC tends to decrease and the average growth of TFP was 4.29% The results also show that women-headed households had higher TE than male-headed ones. Education level and highest degree of household heads, irrigation index and land fragmentation positively and significantly influenced TE while ethnicity of household head negatively influenced TE.

Chapter 4 immerses in analyzing technical inefficiency and production risk of maize farming in the Northestern, Vietnam. This study combines two methods: Just and Pope's stochastic production and a stochastic frontier analysis to analyze panel balance data. By using the likelihood ratio test, the most the optimal model was selected. The findings show that if other inputs are constant, 1% increase in cultivated land, fertilizer utilization, labour, seed or other costs, would increase yields by 0.601%, 0.488%, 0.178%, 0.067%, 0.004%, respectively. If all the inputs are increased by 1%, the output will increase by 0.9%. The estimates results of marginal output risk indicates that land, labour, pesticide and herbicides raise the possibility of output variances, while seed, fertilizer and other costs reduce the output variances. Determinants of technical inefficiency provides gender of the household leaders, the number of household members take positive impact on the technical inefficiency. Conversely, education and highest degree of the household leaders, irrigation, disaster, and land fragmentation have negative impact on technical efficiency. The findings show that maize farmers can boost their production by 15.27% by improving technical efficiency. The average TE of maize production in the North West is 82.75 and increased steadily between 2008 and 2016. Dien Bien province recorded the highest level of TE, followed by Lao Cai and the lowest is Lai Chau.

5.2. Policy implications and further research suggestions

The study findings provide key policy implications on how to improve TE, technical change and return to scale in Vietnam agriculture and in rice and maize production in particular.

The research results of chapter 2 show that all inputs positively influence the value of agricultural production in Vietnam, in which land is the most important factor. Farm size negatively influenced TE. Therefore, it shows that land rotation plays a very important role in improving the output value of agricultural production. Thus, Governments and organizations need to intensify and broaden training of farmers on agricultural production as well as on market access to increase production and efficiency in Vietnam agriculture. This will involve farm diversification into crop and livestock production with improved efficiency under different optimal input combinations to yield desirable output levels. There is disguised unemployment and low labour efficiency in Vietnam. The government needs to facilitate vocational training and career orientation to improve labour quality and also connect the unemployed to the labour market. From a macro economics perspective, Government can use two major policies to create jobs and increase farmers' incomes: (1) Fiscal policy: expanding fiscal policy through money supply expansion to create more jobs in the labour market and people can easily access; (2) Income policy: price tools can be used to change the price correlation between agricultural products with prices of other industries. It is also possible to use wage tools through fixed minimum wage.

The results discussed in chapter 3 show that all inputs positively influenced the value of rice production, in which land was the most important factor. Therefore, two important solutions to solve land fragmentation problem: there should be policies (or legislation) and incentive mechanisms to facilitate land consolidation to increase the cultivated land area. Besides, irrigation has a positive impact on technical efficiency. Therefore, it is necessary to have appropriate investment policies for irrigation systems in some areas with unfavorable natural conditions. Education and the highest level of household leaders variables take positive impact on technical efficiency. Thus, appropriate farmer trainings and robust extension services would be vital to scale up adoption of productivity-increasing technologies in rice production.

Chapter 4 results also indicate that all inputs positively and significantly influence the value of maize production, in which land was the most important factor, followed by labour. Thus, in order to reduce technological inefficiencies in maize production, farmers should strive to further improve their education levels while actively seeking short-term training courses in maize production. Farmers should actively prevent the impact of disease and the external environment on maize production, and optimum inputs for maize production under limited resources. The government should have appropriate investment policies for irrigation systems in areas with unfavorable natural conditions, appropriate land policies and good legal corridors to facilitate the consolidation of plots, avoiding land fragmentation. It is necessary to expand the training courses and apply new technologies to production through various channels such as the Women's Union, the Farmers' Association, the non-governmental organizations and so on.

The study had some methodological limitations as follows:

In analyzing the drivers of TE, Battese and Coelli (1995) model to express (in)efficiency component (U_i) as a function of an exogenous variable and a random error term. However, it was not possible on stata software. Therefore, this study used a second-stage regression by regressing the TE results from the first stage regression. The results on drivers of TE may be erroneous because this method may violate the assumption of independence of TE estimations. An application on Vuong test for non-nested model would have improved reliability of the estimates. For the discussion in chapter 2, the research findings would have been more meaningful if TE had been studied under a certain policy.

This study provides diverse starting points for future researches. Subsequent studies may include more elements of policies that influence TE, output elasticity and return to scale. This study could be a good basis for a detailed study on household resources such as land, labour, capital, seeds, fertilizers, pesticides and herbicide, among other inputs. Furthermore, assessing the role of institutions and market innovations to overcome production and marketing risks would have been insightful, given the contract farming arrangements that exist in Vietnam. Hence, in the future, it would be possible and necessary to research the role of institutions such as farmers associations and contract farming in improving TE, value of production, technical change and returns to scale in Vietnam agriculture.

In general, this dissertation demonstrates the diversity in measuring TE of the agricultural sector in Vietnam in general and in specific agricultural commodities, maize and rice from 2008 to 2016 and provides suggestions to improve TE in the household level.

References

Abdul-Majid, M., Saal, D. S., & Battisti, G. (2008). *The efficiency and productivity of Malaysian banks: an output distance function approach*: Aston University.

Addai, K. N., & Owusu, V. (2014). Technical efficiency of maize farmers across various agro ecological zones of Ghana. *Journal of Agriculture and Environmental Sciences, 3*(1), 149-172.

Ahmed, M. H. (2014). Technical efficiency of maize producing farmers in Arsi Negelle, central rift valley of Ethiopia: stochastic frontier approach. *Agriculture & Forestry/Poljoprivreda i Sumarstvo, 60*(1).

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics, 6*(1), 21-37.

Akter, S. (2003). *Competitiveness and efficiency in poultry and pig production in Vietnam* (Vol. 57): ILRI (aka ILCA and ILRAD).

Anupama, J., Singh, R., & Kumar, R. (2005). Technical efficiency in maize production in Madhya Pradesh: Estimation and implications. *Agricultural Economics Research Review, 18*(2), 305-315.

Aye, G. C., & Mungatana, E. D. (2010). Technical efficiency of traditional and hybrid maize farmers in Nigeria: Comparison of alternative approaches. *African Journal of Agricultural Research, 5*(21), 2909-2917.

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of econometrics, 38*(3), 387-399.

Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of productivity analysis, 3*(1-2), 153-169.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical economics, 20*(2), 325-332.

Battese, G. E., & Corra, G. S. (1977). Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian journal of agricultural economics, 21*(3), 169-179.

Battese, G. E., Rambaldi, A., & Wan, G. (1997). A stochastic frontier production function with flexible risk properties. *Journal of productivity analysis, 8*(3), 269-280.

Bokusheva, R., & Hockmann, H. (2006). Production risk and technical inefficiency in Russian agriculture. *European Review of Agricultural Economics, 33*(1), 93-118.

Brümmer, B., Glauben, T., & Lu, W. (2006). Policy reform and productivity change in Chinese agriculture: A distance function approach. *Journal of Development Economics, 81*(1), 61-79.

Brümmer, B., Glauben, T., & Thijssen, G. (2002). Decomposition of productivity growth using distance functions: The case of dairy farms in three European countries. *American Journal of Agricultural Economics, 84*(3), 628-644.

Chirwa, E. W. (2007). Sources of technical efficiency among smallholder maize farmers in Southern Malawi.

Coelli, T., & Perelman, S. (1996). Efficiency measurement, multiple-output technologie and distance functions: With application to European Railways: Crepp.

Coelli, T., & Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: With application to European railways. *European journal of operational research, 117*(2), 326-339.

Coelli, T., & Perelman, S. (2000). Technical efficiency of European railways: a distance function approach. *Applied Economics, 32*(15), 1967-1976.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*: Springer Science & Business Media.

Cornwell, C., Schmidt, P., & Sickles, R. C. (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of econometrics, 46*(1-2), 185-200.

Cuesta, R. A., & Orea, L. (2002). Mergers and technical efficiency in Spanish savings banks: A stochastic distance function approach. *Journal of Banking & Finance, 26*(12), 2231-2247.

Dang, N. H. (2017). Estimation of Technical Efficiency and Its Determinants of White Maize Production in Vinh Long Province: A Stochastic Production Frontier Approach. *Review of Integrative Business and Economics Research, 6*(4), 341.

Dao, G., & Lewis, P. (2013). Technical efficiency of annual crop farms in northern Vietnam: Working Paper, available at: www. murdoch. edu. au.

De Janvry, A., & Sadoulet, E. (2009). Agricultural growth and poverty reduction: Additional evidence. *The World Bank Research Observer, 25*(1), 1-20.

Den, D. T., Ancev, T., & Harris, M. (2007). Technical efficiency of prawn farms in the Mekong Delta, Vietnam.

Devendra, C. (1999). Goats: Challenges for increased productivity and improved livelihoods. *Outlook on Agriculture, 28*(4), 215-226.

Dinh Thao, T., Tri Khiem, N., Xuan Trieu, M., Gerpacio, R., & Pingali, P. (2004). *Maize in Vietnam: Production systems, constraints, and research priorities*: CIMMYT.

Dlamini, S. I., Masuku, M. B., & Rugambisa, J. (2012). Technical efficiency of maize production in Swaziland: A stochastic frontier approach. *African Journal of Agricultural Research, 7*(42), 5628-5636.

Emvalomatis, G. (2012). Productivity growth in German dairy farming using a flexible modelling approach. *Journal of Agricultural Economics, 63*(1), 83-101.

Essilfie, F. L., Asiamah, M. T., & Nimoh, F. (2011). Estimation of farm level technical efficiency in small scale maize production in the Mfantseman Municipality in the Central Region of Ghana: A stochastic frontier approach. *Journal of Development and Agricultural Economics, 3*(14), 645-654.

Feng, G., & Serletis, A. (2010). Efficiency, technical change, and returns to scale in large US banks: Panel data evidence from an output distance function satisfying theoretical regularity. *Journal of Banking & Finance, 34*(1), 127-138.

Ghulam, S., Farman, A., Khalid, M., Musawar, S., & Inamullah, K. (2009). Technical efficiency of maize farmers in various ecological zones of AJK. *Sarhad Journal of Agriculture, 25*(4), 607-610.

Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of productivity analysis, 23*(1), 7-32.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of econometrics, 126*(2), 269-303.

GSO. (2000). Statistic Yearbook of Vietnam [Online], Hanoi.

GSO. (2011). Statistic Yearbook of Vietnam [Online], Hanoi.

GSO. (2016). Statistic Yearbook of Vietnam [Online], Hanoi.

Ha, D. T. T. (2012). Production Survey of Paddy in Vietnam. *Ministry of Planning and Investment*.

Hetemäki, L. (1996). *Essays on the impact of pollution control on a firm: a distance function approach*: Helsinki Research Centre Helsinki.

Hien, N. T. M., Kawaguchi, T., & Suzuki, N. (2003). A study on technical efficiency of rice production in the Mekong Delta, Vietnam by stochastic frontier analysis. *Journal of the Faculty of Agriculture, 48*(1/2), 325-357.

Hoang Linh, V. (2012). Efficiency of rice farming households in Vietnam. *International Journal of Development Issues, 11*(1), 60-73.

Huynh-Truong, H. (2009). Technical efficiency of rice producing households in the Mekong delta of Vietnam.

Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). Agricultural productivity growth and poverty alleviation. *Development policy review, 19*(4), 449-466.

Isaac, O. (2011). Technical efficiency of maize production in Oyo State. *Journal of Economics and International finance, 3*(4), 211-216.

Jabbar, M. A., & Akter, S. (2006). Market and other factors affecting farm-specific production efficiency in poultry production in Vietnam. *Journal of Food Products Marketing, 12*(2), 99-113.

Jaenicke, E. C., Frechette, D. L., & Larson, J. A. (2003). Estimating production risk and inefficiency simultaneously: An application to cotton cropping systems. *Journal of Agricultural and Resource Economics*, 540-557.

Jiang, C., Yao, S., & Zhang, Z. (2009). The effects of governance changes on bank efficiency in China: A stochastic distance function approach. *China Economic Review, 20*(4), 717-731.

Joshi, P. K., Gulati, A., Birthal, P. S., & Tewari, L. (2004). Agriculture diversification in South Asia: patterns, determinants and policy implications. *Economic and political weekly*, 2457-2467.

Kalirajan, K. P., & Shand, R. T. (1999). Frontier production functions and technical efficiency measures. *Journal of Economic surveys, 13*(2), 149-172.

Kassie, M., Shiferaw, B., & Muricho, G. (2011). Agricultural technology, crop income, and poverty alleviation in Uganda. *World Development, 39*(10), 1784-1795.

Keil, A., Saint-Macary, C., & Zeller, M. (2009). Maize boom in the uplands of Northern Vietnam: economic importance and environmental implications.

Khai, H. V., & Yabe, M. (2011). Technical efficiency analysis of rice production in Vietnam. *Journal of ISSAAS, 17*(1), 135-146.

Khai, H. V., Yabe, M., Yokogawa, H., & Sato, G. (2008). Analysis of productive efficiency of soybean production in the Mekong River Delta of Viet Nam. *JOURNAL-FACULTY OF AGRICULTURE KYUSHU UNIVERSITY, 53*(1), 271.

Khush, G. S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. *Plant molecular biology, 59*(1), 1-6.

Kibaara, B., & Kavoi, M. (2012). Application of stochastic frontier approach model to assess technical efficiency in Kenya's maize production. *Journal of agriculture, science and technology, 14*(1).

Kirchmann, H., & Thorvaldsson, G. (2000). Challenging targets for future agriculture. *European Journal of Agronomy, 12*(3-4), 145-161.

Kompas, T. (2002). Market reform, productivity and efficiency in Vietnamese rice production.

Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of econometrics, 46*(1-2), 201-211.

Kumbhakar, S. C. (2002). Specification and estimation of production risk, risk preferences and technical efficiency. *American Journal of Agricultural Economics, 84*(1), 8-22.

Kuwornu, J. K., Amoah, E., & Seini, W. (2013). Technical efficiency analysis of maize farmers in the Eastern Region of Ghana. *Journal of Social and Development Sciences, 4*(2), 84.

Le, T. L., Pai Po Lee, Ke Chung Peng, & Chung, R. H. (2017). Factors Influencing Technical Efficiency of Rice Farms in Dong Thap Province, Vietnam: An Application of Two-Stage DEA. *American-Eurasian J. Agric. & Environ. Sci., 17*(3), 245-249. doi: 10.5829/idosi.aejaes.2017.245.249

Lemaire, G., Franzluebbers, A., de Faccio Carvalho, P. C., & Dedieu, B. (2014). Integrated crop–livestock systems: Strategies to achieve synergy between agricultural production and environmental quality. *Agriculture, Ecosystems & Environment, 190*, 4-8.

Ligeon, C., Jolly, C., Bencheva, N., Delikostadinov, S., & Puppala, N. (2013). Production efficiency and risks in limited resource farming: The case of Bulgarian peanut industry. *Journal of Development and Agricultural Economics, 5*(4), 150-160.

Lovell, C. K., Travers, P., Richardson, S., & Wood, L. (1994). Resources and functionings: a new view of inequality in Australia *Models and measurement of welfare and inequality* (pp. 787-807): Springer.

Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de estadística, 4*(2), 209-242.

Mango, N., Makate, C., Hanyani-Mlambo, B., Siziba, S., & Lundy, M. (2015). A stochastic frontier analysis of technical efficiency in smallholder maize production in Zimbabwe: The post-fast-track land reform outlook. *Cogent Economics & Finance, 3*(1), 1117189.

Meeusen, W., & van den Broeck, J. (1977). Technical efficiency and dimension of the firm: Some results on the use of frontier production functions. *Empirical economics, 2*(2), 109-122.

Newman, C., & Matthews, A. (2006). The productivity performance of Irish dairy farms 1984–2000: a multiple output distance function approach. *Journal of productivity analysis, 26*(2), 191-205.

Newman, C., & Matthews, A. (2007). Evaluating the productivity performance of agricultural enterprises in Ireland using a multiple output distance function approach. *Journal of Agricultural Economics, 58*(1), 128-151.

O'Donnell, C. J., & Coelli, T. J. (2005). A Bayesian approach to imposing curvature on distance functions. *Journal of econometrics, 126*(2), 493-523.

Ogundari, K., & Akinbogun, O. O. (2010). Modeling technical efficiency with production risk: A study of fish farms in Nigeria. *Marine Resource Economics, 25*(3), 295-308.

Ogundari, K., & Brümmer, B. (2011). Estimating Technical Efficiency, Input substitution and complementary effects using Output Distance Function: A study of Cassava production in Nigeria. *Agricultural Economics Review, 12*(2), 62-79.

Oppong, B. A., Onumah, E. E., & Asuming-Brempong, S. (2016). Technical efficiency and production risk of maize production: evidence from Ghana. *Asian J. Agric. Extens. Econ. Soc., 11*(3), 1-9.

Pedroso, R., Tran, D. H., Viet, T. Q., Le, A. V., Dang, K. T., & Le, K. P. (2018). Technical efficiency of rice production in the delta of the Vu Gia Thu Bon river basin, Central Vietnam. *World Development Perspectives, 9*, 18-26. doi: 10.1016/j.wdp.2017.12.001

Pham, T. T. (2016). Specialty Rice Adoption, Collective Action and Marketing Channel Choice: Insights from Vietnam. *Dissertation*.

Pingali, P. L., & Rosegrant, M. W. (1995). Agricultural commercialization and diversification: processes and policies. *Food policy, 20*(3), 171-185.

PRB. (2017). Population Reference Bureau, Data Finder. http://www.prb.org. Accessed 06 August 2018.

Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences, 1312*(1), 105-112.

Ryan, J. G., & Spencer, D. (2001). *Future challenges and opportunities for agricultural R&D in the semi-arid tropics*: International Crops Research Institute for the Semi-Arid Tropics.

Schmidt, P., & Lovell, C. K. (1979). Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of econometrics, 9*(3), 343-366.

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics, 2*(4), 367-374.

Shapiro, K. H., & Müller, J. (1977). Sources of technical efficiency: the roles of modernization and information. *Economic Development and Cultural Change, 25*(2), 293-310.

Shar, A. M., Mahesar, A. A., Chandio, A. D., & Memon, K. R. (2017). Impact of confining stress on permeability of tight gas sands: an experimental study. *Journal of Petroleum Exploration and Production Technology, 7*(3), 717-726.

Shephard, R. W. (2012). *Cost and production functions* (Vol. 194): Springer Science & Business Media.

Thirtle, C., Lin, L., & Piesse, J. (2003). The impact of research-led agricultural productivity growth on poverty reduction in Africa, Asia and Latin America. *World Development, 31*(12), 1959-1975.

Tiedemann, T., & Latacz-Lohmann, U. (2013). Production risk and technical efficiency in organic and conventional agriculture–the case of arable farms in Germany. *Journal of Agricultural Economics, 64*(1), 73-96.

Tuan M.Cao, Sutonya Thongra, & Kiatpathomchai, S. (2017). Technical Efficiency in Rice Production of the Farmers in Cooperatives: A Study in KienGiang Province, Mekong Delta, Vietnam. *International Journal of Economic Research, 14*(15), 689-704.

Tung, P. B. V. (2010). *Technical efficiency of improved extensive shrimp farming in Ca Mau province, Vietnam.* Universitetet i Tromsø.

Tveteros, R. (1999). Production risk and productivity growth: Some findings for Norwegian salmon aquaculture. *Journal of productivity analysis, 12*(2), 161-179.

Van Hung, P., MacAulay, T. G., & Marsh, S. P. (2007). The economics of land fragmentation in the north of Vietnam. *Australian journal of agricultural and resource economics, 51*(2), 195-211.

Villano, R., & Fleming, E. (2006). Technical inefficiency and production risk in rice farming: evidence from Central Luzon Philippines. *Asian economic journal, 20*(1), 29-46.

Wakili, A. M. (2012). Technical Efficiency of Maize Farmers in Gombi Local Government of Adamawa State, Nigeria. *Agricultural Journal, 7*(1), 1-4.

Wang, C. (2007). Decomposing energy productivity change: a distance function approach. *Energy, 32*(8), 1326-1333.

Appendix

Appendix 2.1. Maximum likelihood estimates of stochastic frontier model (model 2 - Mu assumed to be zero)

Time-varying decay model (truncated-normal)		Number of obs =	2435
Group variable: id		Number of groups =	487
Time variable: yr		Obs per group: min =	5
		avg =	5.0
		max =	5
		Prob > chi2 =	0.0000
Log likelihood = -894.8523		Wald chi2(20) =	15003.08

(1) [Mu]_cons = 0

llivestockneg	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Frontier						
lcropolivestock	1.409238	.1239568	11.37	0.000	1.166287	1.652189
lcropolivestock_sq	-.0877279	.0138823	-6.32	0.000	-.1149366	-.0605191
lland	.1238926	.0697783	1.78	0.076	-.0128702	.2606555
llabor	.5857229	.1819048	3.22	0.001	.229196	.9422497
linter	-.0338938	.126558	-0.27	0.789	-.2819429	.2141553
lland_sq	-.0741029	.0030709	-24.13	0.000	-.0801217	-.0680841
llabor_sq	-.0755605	.0242472	-3.12	0.002	-.1230842	-.0280368
linter_sq	.0160403	.0141258	1.14	0.256	-.0116458	.0437264
llandllabor	.0508425	.0104422	4.87	0.000	.0303762	.0713088
llandlinter	-.0035181	.0065722	-0.54	0.592	-.0163994	.0093633
llaborlinter	-.0645881	.0155514	-4.15	0.000	-.0950682	-.0341079
lcropolivestocklland	.0017663	.0067775	0.26	0.794	-.0115174	.01505
lcropolivestockllabor	-.0687841	.017361	-3.96	0.000	-.102811	-.0347572
lcropolivestocklinter	.0199214	.0113832	1.75	0.080	-.0023893	.0422321
yr	.446707	.0874865	5.11	0.000	.2752366	.6181774
yr2	-.0152697	.0103302	-1.48	0.139	-.0355165	.0049771
yrlcropolivestock	.0024861	.0082471	0.30	0.763	-.0136779	.0186501
yrlland	-.020009	.0059615	-3.36	0.001	-.0316934	-.0083247
yrllabor	-.0312084	.0130332	-2.39	0.017	-.056753	-.0056637
yrlinter	-.0010748	.0078591	-0.14	0.891	-.0164783	.0143287
_cons	-9.322585	.8263728	-11.28	0.000	-10.94225	-7.702924
/lnsigma2	-1.647492	.1007873	-16.35	0.000	-1.845031	-1.449952
/ilgtgamma	-.2698705	.243696	-1.11	0.268	-.7475058	.2077648
/mu	(omitted)					
/eta	-.1741362	.0681296	-2.56	0.011	-.3076678	-.0406046
sigma2	.1925323	.0194048			.1580204	.2345815
gamma	.4329389	.059828			.321365	.5517552
sigma_u2	.0833547	.0196906			.0447619	.1219475
sigma_v2	.1091776	.0034878			.1023416	.1160136

Appendix 2.2. Maximum likelihood estimates of stochastic frontier model (model 3 - Eta assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =      2435
Group variable: id                                Number of groups =       487
Time variable: yr                               Obs per group: min =         5
                                                               avg =       5.0
                                                               max =         5

                                                     Prob > chi2   =    0.0000
Log likelihood =  -898.1312                          Wald chi2(20)  =  14728.74

 ( 1)  [Eta]_cons = 0
------------------------------------------------------------------------------------------
        llivestockneg |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------------+-------------------------------------------------------------------
Frontier              |
      lcropolivestock |   1.422712   .1241957    11.46   0.000     1.179293    1.666131
   lcropolivestock_sq |  -.0898385   .0141046    -6.37   0.000     -.117483    -.062194
                lland |   .1186038   .0702835     1.69   0.092    -.0191494    .2563569
               llabor |   .6077427    .181274     3.35   0.001     .2524523    .9630332
               linter |  -.0394778   .1324453    -0.30   0.766    -.2990657    .2201101
             lland_sq |  -.0747589   .0032723   -22.85   0.000    -.0811724   -.0683453
            llabor_sq |  -.0762217      .0239    -3.19   0.001    -.1230649   -.0293786
            linter_sq |   .0166914   .0147894     1.13   0.259    -.0122953     .045678
          llandllabor |   .0494713   .0105867     4.67   0.000     .0287217    .0702209
          llandlinter |  -.0015663   .0065912    -0.24   0.812    -.0144849    .0113523
         llaborlinter |  -.0648591   .0156031    -4.16   0.000    -.0954405   -.0342776
 lcropolivestocklland |   .0020526   .0068604     0.30   0.765    -.0113936    .0154988
lcropolivestockllabor |  -.0699583   .0175063    -4.00   0.000    -.1042699   -.0356466
lcropolivestocklinter |   .0200076   .0114566     1.75   0.081     -.002447    .0424621
                   yr |   .4889804   .0855509     5.72   0.000     .3213037     .656657
                  yr2 |  -.0203544   .0098245    -2.07   0.038    -.0396101   -.0010987
    yrlcropolivestock |  -.0000589    .008228    -0.01   0.994    -.0161856    .0160677
              yrlland |  -.0207672   .0059631    -3.48   0.000    -.0324546   -.0090797
             yrllabor |  -.0322149   .0129639    -2.48   0.013    -.0576236   -.0068062
             yrlinter |   -.005435   .0077227    -0.70   0.482    -.0205711    .0097012
                _cons |  -9.387351   .8368451   -11.22   0.000    -11.02754   -7.747164
----------------------+-------------------------------------------------------------------
   /lnsigma2          |   -1.32377   1.491005    -0.89   0.375    -4.246085    1.598545
  /ilgtgamma          |   .3378784   2.538329     0.13   0.894    -4.637154    5.312911
         /mu          |   -.886944    3.16977    -0.28   0.780    -7.099579    5.325691
        /eta          |  (omitted)
-------------+----------------------------------------------------------------------------
      sigma2          |   .2661301   .3968012                      .0143202    4.945833
       gamma          |   .5836751     .61681                      .0095923    .9950966
     sigma_u2         |   .1553335   .3957325                     -.6202879    .9309549
     sigma_v2         |   .1107966   .0036916                      .1035612    .1180319
------------------------------------------------------------------------------------------
```

Appendix 2.3. Maximum likelihood estimates of stochastic frontier model (model 4 - Mu and Eta assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =      2435
Group variable: id                               Number of groups =       487
Time variable: yr                              Obs per group: min =         5
                                                              avg =       5.0
                                                              max =         5

                                                    Prob > chi2   =    0.0000
Log likelihood =  -898.3225                         Wald chi2(20)  =  15045.06

 ( 1)  [Mu]_cons = 0
 ( 2)  [Eta]_cons = 0
--------------------------------------------------------------------------------------
        llivestockneg |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------------+---------------------------------------------------------------
Frontier              |
      lcropolivestock |   1.423243   .1239859    11.48   0.000     1.180235     1.66625
   lcropolivestock_sq |  -.0923966   .0137564    -6.72   0.000    -.1193587   -.0654346
                lland |   .1181758   .0700299     1.69   0.092    -.0190803    .2554319
               llabor |   .6162923   .1810047     3.40   0.001     .2615297    .9710549
               linter |  -.0051657   .1249632    -0.04   0.967    -.2500891    .2397576
             lland_sq |  -.0738833   .0031186   -23.69   0.000    -.0799956    -.067771
            llabor_sq |  -.0744213   .0238473    -3.12   0.002    -.1211612   -.0276813
            linter_sq |   .0129269   .0139972     0.92   0.356    -.0145071    .0403609
          llandllabor |   .0483566   .0104982     4.61   0.000     .0277806    .0689327
          llandlinter |  -.0011417   .0065391    -0.17   0.861    -.0139581    .0116747
         llaborlinter |  -.0662254   .0154611    -4.28   0.000    -.0965286   -.0359223
  lcropolivestocklland |   .0022115   .0068353     0.32   0.746    -.0111855    .0156085
lcropolivestockllabor |  -.0698029   .0174488    -4.00   0.000     -.104002   -.0356039
lcropolivestocklinter |   .0202915   .0112684     1.80   0.072    -.0017941    .0423771
                   yr |   .4832028   .0852172     5.67   0.000     .3161802    .6502254
                  yr2 |  -.0200127   .0097829    -2.05   0.041    -.0391868   -.0008387
    yrlcropolivestock |   .0001848   .0082049     0.02   0.982    -.0158966    .0162662
               yrlland |  -.0212837   .0059402    -3.58   0.000    -.0329262   -.0096411
              yrllabor |  -.0312278   .0129115    -2.42   0.016    -.0565339   -.0059217
              yrlinter |   -.005028   .0077004    -0.65   0.514    -.0201206    .0100645
                _cons |  -9.530643   .8185677   -11.64   0.000    -11.13501   -7.926279
----------------------+---------------------------------------------------------------
   /lnsigma2          |  -1.864903   .0487125   -38.28   0.000    -1.960378   -1.769428
  /ilgtgamma       |  -.8953373   .1847464    -4.85   0.000    -1.257434    -.533241
         /mu       |  (omitted)
        /eta       |  (omitted)
-------------+------------------------------------------------------------------------
      sigma2      |   .1549112   .0075461                          .1408052    .1704304
       gamma      |   .2900096     .03804                           .221416    .3697613
    sigma_u2      |   .0449258   .0077679                          .0297009    .0601506
    sigma_v2      |   .1099855   .0035277                          .1030713    .1168997
--------------------------------------------------------------------------------------
```

Appendix 2.4. Maximum likelihood estimates of stochastic frontier model (model 5 - Mu, Eta and Gamma assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =      2435
Group variable: id                               Number of groups =       487
Time variable: yr                              Obs per group: min =         5
                                                              avg =       5.0
                                                              max =         5

                                                    Prob > chi2   =    0.0000
Log likelihood =  -928.7543                         Wald chi2(20)  =  16147.60

 ( 1)  [Mu]_cons = 0
 ( 2)  [Eta]_cons = 0
 ( 3)  [Gamma]_cons = -100
--------------------------------------------------------------------------------------
        llivestockneg |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------------+---------------------------------------------------------------
Frontier              |
      lcropolivestock |   1.383655   .1251209    11.06   0.000     1.138422    1.628887
   lcropolivestock_sq |  -.0958321    .013794    -6.95   0.000    -.1228679   -.0687964
                lland |   .0888975   .0712184     1.25   0.212     -.050688     .228483
               llabor |    .585754   .1844116     3.18   0.001     .2243139     .947194
               linter |   .1085075   .1215844     0.89   0.372    -.1297935    .3468085
             lland_sq |  -.0760956   .0030016   -25.35   0.000    -.0819786   -.0702126
            llabor_sq |   -.080494   .0242569    -3.32   0.001    -.1280367   -.0329512
            linter_sq |  -.0035948   .0136855    -0.26   0.793    -.0304179    .0232284
          llandllabor |   .0553837   .0106046     5.22   0.000      .034599    .0761684
          llandlinter |   -.002249    .006589    -0.34   0.733    -.0151632    .0106652
         llaborlinter |  -.0649362   .0156379    -4.15   0.000    -.0955859   -.0342864
 lcropolivestocklland |   .0025529   .0069098     0.37   0.712    -.0109902    .0160959
lcropolivestockllabor |  -.0636469   .0176061    -3.62   0.000    -.0981543   -.0291395
lcropolivestocklinter |   .0209387   .0110353     1.90   0.058    -.0006902    .0425675
                   yr |   .4746042   .0882165     5.38   0.000     .3017031    .6475054
                  yr2 |  -.0254626   .0103007    -2.47   0.013    -.0456516   -.0052736
    yrlcropolivestock |   .0000846   .0083356     0.01   0.992     -.016253    .0164222
              yrlland |   -.018755   .0061241    -3.06   0.002     -.030758    -.006752
             yrllabor |  -.0375008   .0132113    -2.84   0.005    -.0633944   -.0116071
             yrlinter |  -.0005613    .007919    -0.07   0.943    -.0160822    .0149597
                _cons |  -9.787007   .8145536   -12.02   0.000     -11.3835   -8.190512
----------------------+---------------------------------------------------------------
   /lnsigma2          |   -2.07504   .0286593   -72.40   0.000    -2.131211   -2.018869
  /ilgtgamma          |       -100          .        .       .            .           .
         /mu          |  (omitted)
        /eta          |  (omitted)
-------------+------------------------------------------------------------------------
      sigma2          |   .1255514   .0035982                      .1186935    .1328056
       gamma          |   3.72e-44          .                             .           .
    sigma_u2          |   4.67e-45          .                             .           .
    sigma_v2          |   .1255514          .                             .           .
--------------------------------------------------------------------------------------
```

Appendix 3.1. Maximum likelihood estimates of stochastic frontier model (model 2 - CD production stochastic frontier form)

```
Time-varying decay model (truncated-normal)        Number of obs   =      7775
Group variable: id1                             Number of groups   =      1555
Time variable: yr                             Obs per group: min   =         5
                                                             avg   =       5.0
                                                             max   =         5

                                                   Prob > chi2     =    0.0000
Log likelihood = -2675.9610                        Wald chi2(12)   =  41755.43

------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .7531194   .0071784   104.91   0.000     .7390499    .7671888
        lnX2 |    .086862   .0074911    11.60   0.000     .0721797    .1015444
        lnX3 |   .0066259   .0021892     3.03   0.002     .0023352    .0109167
        lnX4 |   .0300413   .0023582    12.74   0.000     .0254194    .0346633
        lnX5 |   .0305852   .0026221    11.66   0.000     .0254461    .0357243
        lnX6 |   .0087247   .0023802     3.67   0.000     .0040597    .0133898
          yr |   .0017248   .0048127     0.36   0.720    -.0077079    .0111575
             |
  regi_code1 |
          2  |  -.1400078   .0175718    -7.97   0.000    -.1744479   -.1055677
          3  |  -.2174801   .0140613   -15.47   0.000    -.2450397   -.1899205
          4  |  -.1114778    .020928    -5.33   0.000    -.1524961   -.0704596
          5  |     .05545   .0180787     3.07   0.002     .0200164    .0908837
          6  |   .3902772   .0274049    14.24   0.000     .3365646    .4439899
             |
       _cons |    .570782   .0491895    11.60   0.000     .4743723    .6671918
-------------+----------------------------------------------------------------
   /lnsigma2 |   .2186465   .5267636     0.42   0.678    -.8137912    1.251084
  /ilgtgamma |   2.389052   .5757404     4.15   0.000     1.260622    3.517483
         /mu |  -13.47266   7.754537    -1.74   0.082    -28.67127    1.725958
        /eta |   .1848815   .0305895     6.04   0.000     .1249271    .2448358
-------------+----------------------------------------------------------------
      sigma2 |   1.244391      .6555                      .4431747    3.494129
       gamma |   .9159887   .0443052                      .7791331    .9711811
    sigma_u2 |   1.139848   .6555356                     -.1449779    2.424675
    sigma_v2 |    .104543   .0018336                      .1009491    .1081368
------------------------------------------------------------------------------
```

Appendix 3.2. Maximum likelihood estimates of stochastic frontier model (model 3 - translog production stochastic frontier form with no technical change)

```
Time-varying decay model (truncated-normal)        Number of obs =       7775
Group variable: id1                             Number of groups =       1555
Time variable: yr                             Obs per group: min =          5
                                                             avg =        5.0
                                                             max =          5

                                                Prob > chi2   =     0.0000
Log likelihood = -1934.5364                     Wald chi2(32)  =  54483.31

------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .7896948   .0606417    13.02   0.000     .6708392    .9085503
        lnX2 |   .1668882   .0512469     3.26   0.001     .0664461    .2673303
        lnX3 |  -.0509971   .0185216    -2.75   0.006    -.0872989   -.0146953
        lnX4 |  -.0829361    .020569    -4.03   0.000    -.1232506   -.0426216
        lnX5 |   .0738715   .0227566     3.25   0.001     .0292694    .1184737
        lnX6 |  -.0441945   .0210033    -2.10   0.035    -.0853602   -.0030288
     lnX1_sq |   .0135971   .0101357     1.34   0.180    -.0062686    .0334628
     lnX2_sq |    .052722   .0135931     3.88   0.000       .02608    .0793639
     lnX3_sq |   .0450405   .0021928    20.54   0.000     .0407426    .0493384
     lnX4_sq |   .0302003   .0018505    16.32   0.000     .0265735    .0338272
     lnX5_sq |   .0134512   .0021009     6.40   0.000     .0093335     .017569
     lnX6_sq |   .0239917   .0017364    13.82   0.000     .0205884    .0273949
    lnX1lnX2 |  -.0416042   .0090106    -4.62   0.000    -.0592647   -.0239437
    lnX1lnX3 |  -.0149093    .002632    -5.66   0.000    -.0200679   -.0097506
    lnX1lnX4 |   .0030865   .0031403     0.98   0.326    -.0030684    .0092413
    lnX1lnX5 |  -.0094705   .0035867    -2.64   0.008    -.0165003   -.0024407
    lnX1lnX6 |  -.0045664   .0031516    -1.45   0.147    -.0107435    .0016106
    lnX2lnX3 |   .0072646   .0031059     2.34   0.019     .0011772     .013352
    lnX2lnX4 |  -.0055954   .0035611    -1.57   0.116     -.012575    .0013841
    lnX2lnX5 |  -.0005373   .0042347    -0.13   0.899    -.0088372    .0077625
    lnX2lnX6 |   .0039883   .0035352     1.13   0.259    -.0029406    .0109172
    lnX3lnX4 |   .0015191   .0010484     1.45   0.147    -.0005357    .0035738
    lnX3lnX5 |    .000362   .0013475     0.27   0.788    -.0022791    .0030031
    lnX3lnX6 |  -.0023246   .0008621    -2.70   0.007    -.0040143   -.0006349
    lnX4lnX5 |  -.0013555   .0009514    -1.42   0.154    -.0032203    .0005092
    lnX4lnX6 |  -.0017492   .0010456    -1.67   0.094    -.0037985    .0003001
    lnX5lnX6 |  -.0031133   .0012683    -2.45   0.014    -.0055991   -.0006276
             |
  regi_code1 |
          2  |  -.1874242   .0152689   -12.27   0.000    -.2173507   -.1574978
          3  |  -.2194368   .0130043   -16.87   0.000    -.2449248   -.1939487
          4  |  -.2045287   .0183755   -11.13   0.000     -.240544   -.1685134
          5  |  -.0994503   .0155919    -6.38   0.000    -.1300098   -.0688908
          6  |   .0884168   .0263356     3.36   0.001        .0368    .1400335
             |
       _cons |   .7600277   .2127918     3.57   0.000     .3429635    1.177092
-------------+----------------------------------------------------------------
   /lnsigma2 |  -1.012588   .5935195    -1.71   0.088    -2.175865    .1506891
  /ilgtgamma |   1.147557    .782928     1.47   0.143    -.3869532    2.682068
```

```
         /mu |  -10.74459   8.118346    -1.32   0.186    -26.65626    5.167072
        /eta |   .5045304   .0523056     9.65   0.000     .4020134    .6070475
-------------+----------------------------------------------------------------
      sigma2 |   .3632777   .2156124                        .11351    1.162635
       gamma |   .7590645   .1431862                       .404451    .9359602
    sigma_u2 |   .2757512   .2156632                      -.146941    .6984434
    sigma_v2 |   .0875265     .00152                      .0845473    .0905057
------------------------------------------------------------------------------
```

Appendix 3.3. Maximum likelihood estimates of stochastic frontier model (model 4- Gamma assumed to be zero)

```
Time-varying decay model (truncated-normal)        Number of obs =      7775
Group variable: id1                              Number of groups =      1555
Time variable: yr                               Obs per group: min =         5
                                                               avg =       5.0
                                                               max =         5

                                                   Prob > chi2   =    0.0000
Log likelihood = -2033.2126                        Wald chi2(40)  =  60662.82

 ( 1)  [Frontier]1b.regi_code1 = 0
 ( 2)  [Gamma]_cons = -100
------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .8301874   .0606486    13.69   0.000     .7113183    .9490564
        lnX2 |   .1437437   .0577828     2.49   0.013     .0304915     .256996
        lnX3 |  -.0424874   .0212884    -2.00   0.046    -.0842119    -.000763
        lnX4 |   -.070203   .0209547    -3.35   0.001    -.1112734   -.0291326
        lnX5 |   .0610508   .0240461     2.54   0.011     .0139213    .1081803
        lnX6 |  -.0805531   .0222187    -3.63   0.000     -.124101   -.0370053
          yr |   .0339505   .0349549     0.97   0.331    -.0345599    .1024609
     lnX1_sq |  -.0029431   .0099295    -0.30   0.767    -.0224046    .0165184
     lnX2_sq |   .0401688   .0146272     2.75   0.006     .0115001    .0688376
     lnX3_sq |   .0437014   .0029943    14.60   0.000     .0378328    .0495701
     lnX4_sq |   .0329704   .0019066    17.29   0.000     .0292336    .0367072
     lnX5_sq |   .0159582   .0021779     7.33   0.000     .0116895    .0202269
     lnX6_sq |   .0177124   .0021018     8.43   0.000      .013593    .0218318
       yr_sq |  -.0082729   .0058903    -1.40   0.160    -.0198178    .0032719
    lnX1lnX2 |  -.0285242   .0094039    -3.03   0.002    -.0469555   -.0100929
    lnX1lnX3 |  -.0143592   .0028112    -5.11   0.000     -.019869   -.0088494
    lnX1lnX4 |   .0053686   .0030461     1.76   0.078    -.0006016    .0113388
    lnX1lnX5 |  -.0094456   .0036895    -2.56   0.010    -.0166768   -.0022144
    lnX1lnX6 |  -.0020347   .0031439    -0.65   0.518    -.0081967    .0041273
      lnX1yr |   .0022486   .0052732     0.43   0.670    -.0080866    .0125838
    lnX2lnX3 |   .0057868   .0035054     1.65   0.099    -.0010836    .0126572
    lnX2lnX4 |  -.0120935   .0036809    -3.29   0.001     -.019308   -.0048791
    lnX2lnX5 |   .0016695   .0046998     0.36   0.722     -.007542     .010881
    lnX2lnX6 |   .0089408   .0037447     2.39   0.017     .0016013    .0162803
      lnX2yr |  -.0091014   .0062509    -1.46   0.145     -.021353    .0031501
    lnX3lnX4 |   .0028722   .0010704     2.68   0.007     .0007743      .00497
    lnX3lnX5 |  -.0010375   .0014344    -0.72   0.469    -.0038487    .0017738
    lnX3lnX6 |   -.003611   .0008933    -4.04   0.000    -.0053619   -.0018602
      lnX3yr |   .0033229   .0030597     1.09   0.277    -.0026741    .0093199
    lnX4lnX5 |  -.0022676   .0009806    -2.31   0.021    -.0041896   -.0003456
    lnX4lnX6 |  -.0009294   .0010955    -0.85   0.396    -.0030766    .0012177
      lnX4yr |  -.0065832   .0021688    -3.04   0.002     -.010834   -.0023323
    lnX5lnX6 |  -.0030654   .0013217    -2.32   0.020    -.0056559   -.0004749
      lnX5yr |    .002705   .0022614     1.20   0.232    -.0017272    .0071373
      lnX6yr |   .0100431   .0022635     4.44   0.000     .0056067    .0144795
             |
  regi_code1 |
```

2	-.1888439	.0145738	-12.96	0.000	-.2174081	-.1602797
3	-.2238984	.0127395	-17.58	0.000	-.2488673	-.1989294
4	-.2082579	.0177667	-11.72	0.000	-.24308	-.1734358
5	-.1357074	.0144146	-9.41	0.000	-.1639596	-.1074553
6	.030449	.0239983	1.27	0.205	-.0165868	.0774849
_cons	.512928	.2289718	2.24	0.025	.0641515	.9617046
/lnsigma2	-2.314865	.0160385	-144.33	0.000	-2.346299	-2.28343
/ilgtgamma	-100	.	.	.	.	.
/mu	(omitted)					
/eta	.0958369	.	.	.	.	.
sigma2	.0987796	.0015843			.0957227	.101934
gamma	3.72e-44	.			.	.
sigma_u2	3.67e-45	.			.	.
sigma_v2	.0987796	.			.	.

Appendix 3.4. Maximum likelihood estimates of stochastic frontier model (model 5- Mu assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =      7775
Group variable: id1                               Number of groups =      1555
Time variable: yr                               Obs per group: min =         5
                                                               avg =       5.0
                                                               max =         5

                                                     Prob > chi2   =    0.0000
Log likelihood = -1933.4973                          Wald chi2(40)  =  51652.95

 ( 1)  [Frontier]1b.regi_code1 = 0
 ( 2)  [Mu]_cons = 0
------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .7563222   .0625052    12.10   0.000     .6338143    .8788301
        lnX2 |   .1953689    .057053     3.42   0.001     .0835472    .3071906
        lnX3 |  -.0448772   .0216074    -2.08   0.038    -.0872269   -.0025275
        lnX4 |  -.0672153   .0214553    -3.13   0.002    -.1092669   -.0251638
        lnX5 |   .0615626   .0237509     2.59   0.010     .0150117    .1081135
        lnX6 |  -.0768253   .0224782    -3.42   0.001    -.1208818   -.0327688
          yr |  -.0703138   .0381419    -1.84   0.065    -.1450705     .004443
     lnX1_sq |   .0169161   .0101912     1.66   0.097    -.0030582    .0368905
     lnX2_sq |   .0498578   .0144346     3.45   0.001     .0215665     .078149
     lnX3_sq |   .0424116    .002924    14.50   0.000     .0366807    .0481425
     lnX4_sq |   .0308666   .0018939    16.30   0.000     .0271545    .0345786
     lnX5_sq |   .0141912   .0021246     6.68   0.000     .0100272    .0183553
     lnX6_sq |   .0181283   .0020988     8.64   0.000     .0140148    .0222418
       yr_sq |   .0088893   .0070702     1.26   0.209     -.004968    .0227466
    lnX1lnX2 |   -.043049   .0094504    -4.56   0.000    -.0615715   -.0245266
    lnX1lnX3 |  -.0162046   .0028422    -5.70   0.000    -.0217752    -.010634
    lnX1lnX4 |   .0030913   .0031449     0.98   0.326    -.0030726    .0092553
    lnX1lnX5 |  -.0092516   .0036755    -2.52   0.012    -.0164556   -.0020477
    lnX1lnX6 |  -.0024092   .0032162    -0.75   0.454    -.0087128    .0038944
      lnX1yr |   .0007737   .0052099     0.15   0.882    -.0094376     .010985
    lnX2lnX3 |   .0099037   .0035968     2.75   0.006      .002854    .0169534
    lnX2lnX4 |  -.0080998   .0036766    -2.20   0.028    -.0153058   -.0008939
    lnX2lnX5 |   .0003149   .0046071     0.07   0.946    -.0087148    .0093445
    lnX2lnX6 |   .0085779   .0038121     2.25   0.024     .0011064    .0160494
      lnX2yr |  -.0108186   .0061351    -1.76   0.078    -.0228431     .001206
    lnX3lnX4 |   .0022274   .0010986     2.03   0.043     .0000742    .0043805
    lnX3lnX5 |   -.000223   .0014617    -0.15   0.879    -.0030879    .0026419
    lnX3lnX6 |  -.0035336   .0009287    -3.80   0.000    -.0053538   -.0017134
      lnX3yr |   .0054852   .0029839     1.84   0.066    -.0003631    .0113335
    lnX4lnX5 |  -.0017109   .0009644    -1.77   0.076    -.0036011    .0001793
    lnX4lnX6 |  -.0013934   .0011101    -1.26   0.209    -.0035692    .0007824
      lnX4yr |  -.0041909   .0021293    -1.97   0.049    -.0083643   -.0000176
    lnX5lnX6 |  -.0030197   .0013205    -2.29   0.022    -.0056078   -.0004316
      lnX5yr |    .002625   .0022078     1.19   0.234    -.0017022    .0069521
```

```
      lnX6yr |   .0099009   .0022303     4.44   0.000     .0055297    .0142722
             |
  regi_code1 |
          2  |    -.18799    .016019   -11.74   0.000    -.2193866   -.1565934
          3  |  -.2285047   .0135684   -16.84   0.000    -.2550983   -.2019111
          4  |  -.2144308   .0192592   -11.13   0.000    -.2521782   -.1766834
          5  |  -.0999517   .0160789    -6.22   0.000    -.1314658   -.0684376
          6  |   .0908918   .0266819     3.41   0.001     .0385962    .1431873
             |
       _cons |   1.054282   .2390381     4.41   0.000      .585776    1.522788
-------------+----------------------------------------------------------------
   /lnsigma2 |  -2.409919   .0216889  -111.11   0.000    -2.452429    -2.36741
  /ilgtgamma |  -3.367735   .4429693    -7.60   0.000    -4.235939   -2.499531
         /mu |  (omitted)
        /eta |   .4339249   .0635545     6.83   0.000     .3093605    .5584894
-------------+----------------------------------------------------------------
      sigma2 |   .0898226   .0019482                      .0860843    .0937232
       gamma |   .0333192   .0142676                      .0142599    .0758911
    sigma_u2 |   .0029928   .0013206                      .0004044    .0055812
    sigma_v2 |   .0868297   .0015367                      .0838179    .0898415
------------------------------------------------------------------------------
```

Appendix 3.5. Maximum likelihood estimates of stochastic frontier model (model 6- Eta assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =       7775
Group variable: id1                               Number of groups =       1555
Time variable: yr                               Obs per group: min =          5
                                                               avg =        5.0
                                                               max =          5

                                                    Prob > chi2   =     0.0000
Log likelihood = -1958.1141                         Wald chi2(40)  =  50976.28

 ( 1)  [Frontier]1b.regi_code1 = 0
 ( 2)  [Eta]_cons = 0
------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .7939129   .0618223    12.84   0.000     .6727433    .9150825
        lnX2 |   .1655763    .057058     2.90   0.004     .0537448    .2774079
        lnX3 |  -.0411479    .020826    -1.98   0.048     -.081966   -.0003297
        lnX4 |   -.068164   .0211161    -3.23   0.001    -.1095508   -.0267771
        lnX5 |   .0570817   .0238045     2.40   0.016     .0104257    .1037377
        lnX6 |  -.0803556   .0219772    -3.66   0.000      -.12343   -.0372812
          yr |   .0399219   .0338826     1.18   0.239    -.0264868    .1063306
     lnX1_sq |   .0063096    .009974     0.63   0.527    -.0132391    .0258582
     lnX2_sq |   .0425511   .0143564     2.96   0.003     .0144132     .070689
     lnX3_sq |   .0420107   .0029547    14.22   0.000     .0362195    .0478018
     lnX4_sq |   .0317661    .001901    16.71   0.000     .0280402    .0354919
     lnX5_sq |   .0149477    .002143     6.98   0.000     .0107476    .0191479
     lnX6_sq |   .0176828   .0020738     8.53   0.000     .0136181    .0217474
       yr_sq |  -.0095721   .0056614    -1.69   0.091    -.0206683     .001524
    lnX1lnX2 |  -.0344757   .0092917    -3.71   0.000    -.0526871   -.0162643
    lnX1lnX3 |  -.0147552    .002748    -5.37   0.000    -.0201412   -.0093692
    lnX1lnX4 |   .0032648   .0030398     1.07   0.283    -.0026931    .0092227
    lnX1lnX5 |  -.0088061    .003653    -2.41   0.016    -.0159659   -.0016462
    lnX1lnX6 |  -.0014987   .0030945    -0.48   0.628    -.0075638    .0045665
      lnX1yr |   .0017103   .0051526     0.33   0.740    -.0083886    .0118093
    lnX2lnX3 |   .0069662   .0034272     2.03   0.042     .0002489    .0136834
    lnX2lnX4 |  -.0085064    .003629    -2.34   0.019     -.015619   -.0013938
    lnX2lnX5 |   .0010187   .0046252     0.22   0.826    -.0080465    .0100838
    lnX2lnX6 |   .0085216   .0036809     2.32   0.021     .0013072     .015736
      lnX2yr |  -.0102615    .006129    -1.67   0.094     -.022274     .001751
    lnX3lnX4 |   .0027065   .0010491     2.58   0.010     .0006503    .0047627
    lnX3lnX5 |  -.0007424   .0014065    -0.53   0.598    -.0034992    .0020144
    lnX3lnX6 |  -.0036262   .0008752    -4.14   0.000    -.0053415    -.001911
      lnX3yr |   .0047144   .0030108     1.57   0.117    -.0011867    .0106155
    lnX4lnX5 |  -.0019558   .0009677    -2.02   0.043    -.0038525   -.0000592
    lnX4lnX6 |  -.0011032   .0010788    -1.02   0.306    -.0032177    .0010113
      lnX4yr |  -.0065541   .0021257    -3.08   0.002    -.0107204   -.0023878
    lnX5lnX6 |  -.0034718   .0013008    -2.67   0.008    -.0060214   -.0009222
```

lnX5yr	.0033503	.0022166	1.51	0.131	-.0009941	.0076947
lnX6yr	.0097684	.0022174	4.41	0.000	.0054225	.0141144
regi_code1						
2	-.1850653	.0161946	-11.43	0.000	-.2168063	-.1533244
3	-.2212728	.013818	-16.01	0.000	-.2483556	-.19419
4	-.1988213	.0196388	-10.12	0.000	-.2373126	-.16033
5	-.1034255	.0164154	-6.30	0.000	-.135599	-.0712519
6	.06321	.0268717	2.35	0.019	.0105425	.1158774
_cons	.7159123	.2380772	3.01	0.003	.2492896	1.182535
/lnsigma2	.235113	.5837248	0.40	0.687	-.9089665	1.379193
/ilgtgamma	2.576689	.6286819	4.10	0.000	1.344495	3.808883
/mu	-11.80857	7.557855	-1.56	0.118	-26.6217	3.00455
/eta	(omitted)					
sigma2	1.265052	.738442			.4029405	3.971693
gamma	.9293462	.0412804			.7932282	.978308
sigma_u2	1.175671	.7384684			-.2717005	2.623043
sigma_v2	.0893807	.0015619			.0863194	.092442

Appendix 3.6. Maximum likelihood estimates of stochastic frontier model (model 7- Mu and Eta assumed to be zero)

```
Time-varying decay model (truncated-normal)         Number of obs =       7775
Group variable: id1                              Number of groups =       1555
Time variable: yr                              Obs per group: min =          5
                                                              avg =        5.0
                                                              max =          5

                                                    Prob > chi2   =     0.0000
Log likelihood = -1968.3198                         Wald chi2(40)  =  48001.90

 ( 1)  [Frontier]1b.regi_code1 = 0
 ( 2)  [Mu]_cons = 0
 ( 3)  [Eta]_cons = 0
```

lnquan_out~t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Frontier						
lnX1	.7843615	.0623198	12.59	0.000	.662217	.9065061
lnX2	.1656463	.0573067	2.89	0.004	.0533272	.2779654
lnX3	-.0398162	.0208648	-1.91	0.056	-.0807104	.0010779
lnX4	-.0654259	.0211087	-3.10	0.002	-.1067981	-.0240537
lnX5	.0576942	.0238678	2.42	0.016	.0109142	.1044743
lnX6	-.081627	.0220164	-3.71	0.000	-.1247784	-.0384756
yr	.0382533	.0339001	1.13	0.259	-.0281898	.1046963
lnX1_sq	.0064706	.0099899	0.65	0.517	-.0131091	.0260504
lnX2_sq	.0418555	.0143745	2.91	0.004	.0136819	.070029
lnX3_sq	.0425198	.0029576	14.38	0.000	.036723	.0483167
lnX4_sq	.0317308	.0019053	16.65	0.000	.0279964	.0354652
lnX5_sq	.0148739	.0021507	6.92	0.000	.0106587	.0190891
lnX6_sq	.0176215	.0020794	8.47	0.000	.013546	.021697
yr_sq	-.0095895	.0056628	-1.69	0.090	-.0206884	.0015093
lnX1lnX2	-.0334558	.0092924	-3.60	0.000	-.0516685	-.015243
lnX1lnX3	-.0150034	.0027509	-5.45	0.000	-.0203951	-.0096117
lnX1lnX4	.0033514	.0030434	1.10	0.271	-.0026136	.0093165
lnX1lnX5	-.0087525	.0036561	-2.39	0.017	-.0159184	-.0015867
lnX1lnX6	-.0014827	.0030976	-0.48	0.632	-.007554	.0045886
lnX1yr	.0021967	.005155	0.43	0.670	-.0079069	.0123004
lnX2lnX3	.006837	.0034336	1.99	0.046	.0001074	.0135667
lnX2lnX4	-.0091617	.0036306	-2.52	0.012	-.0162775	-.0020458
lnX2lnX5	.0009475	.0046284	0.20	0.838	-.0081241	.010019
lnX2lnX6	.0086806	.0036863	2.35	0.019	.0014556	.0159056
lnX2yr	-.010601	.0061356	-1.73	0.084	-.0226266	.0014247
lnX3lnX4	.002669	.0010508	2.54	0.011	.0006094	.0047286
lnX3lnX5	-.0006939	.0014097	-0.49	0.623	-.0034568	.002069
lnX3lnX6	-.0035893	.0008771	-4.09	0.000	-.0053083	-.0018702
lnX3yr	.0043836	.0030128	1.45	0.146	-.0015215	.0102886
lnX4lnX5	-.0019394	.0009687	-2.00	0.045	-.0038381	-.0000407
lnX4lnX6	-.001057	.0010801	-0.98	0.328	-.003174	.00106
lnX4yr	-.0064564	.0021242	-3.04	0.002	-.0106197	-.0022931

```
   lnX5lnX6 |  -.0034851   .0013023    -2.68   0.007    -.0060376   -.0009325
     lnX5yr |   .0031438   .0022188     1.42   0.157     -.001205    .0074926
     lnX6yr |   .0099603   .0022213     4.48   0.000     .0056067    .0143139
            |
 regi_code1 |
          2 |  -.1863861   .0167793   -11.11   0.000     -.219273   -.1534993
          3 |  -.2238524   .0142061   -15.76   0.000    -.2516959   -.1960089
          4 |  -.2030689   .0202516   -10.03   0.000    -.2427613   -.1633765
          5 |  -.1126299   .0166673    -6.76   0.000    -.1452973   -.0799625
          6 |   .0613084   .0271948     2.25   0.024     .0080076    .1146092
            |
      _cons |   .7874786   .2406179     3.27   0.001     .3158761    1.259081
-------------+----------------------------------------------------------------
  /lnsigma2 |  -2.157987   .0246508   -87.54   0.000    -2.206302   -2.109673
 /ilgtgamma |  -1.219973   .1145822   -10.65   0.000     -1.44455   -.9953962
        /mu |  (omitted)
       /eta |  (omitted)
-------------+----------------------------------------------------------------
     sigma2 |   .1155574   .0028486                       .110107    .1212776
      gamma |   .2279412   .0201646                      .1908417    .2698475
   sigma_u2 |   .0263403   .0028539                      .0207469    .0319337
   sigma_v2 |   .0892171   .0015809                      .0861185    .0923158
------------------------------------------------------------------------------
```

Appendix 3.7. Maximum likelihood estimates of stochastic frontier model (model 8- Gamma, Mu and Eta assumed to be zero)

```
Time-varying decay model (truncated-normal)        Number of obs =      7775
Group variable: id1                              Number of groups =      1555
Time variable: yr                              Obs per group: min =         5
                                                              avg =       5.0
                                                              max =         5

                                                   Prob > chi2   =    0.0000
Log likelihood = -2033.2126                        Wald chi2(40)  =  60662.80

 ( 1)  [Frontier]1b.regi_code1 = 0
 ( 2)  [Mu]_cons = 0
 ( 3)  [Eta]_cons = 0
 ( 4)  [Gamma]_cons = -100
------------------------------------------------------------------------------
lnquan_out~t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Frontier     |
        lnX1 |   .8301874   .0606486    13.69   0.000     .7113183    .9490564
        lnX2 |   .1437437   .0577828     2.49   0.013     .0304915     .256996
        lnX3 |  -.0424874   .0212884    -2.00   0.046    -.0842119    -.000763
        lnX4 |   -.070203   .0209547    -3.35   0.001    -.1112734   -.0291325
        lnX5 |   .0610508   .0240461     2.54   0.011     .0139213    .1081803
        lnX6 |  -.0805531   .0222187    -3.63   0.000     -.124101   -.0370053
          yr |   .0339505   .0349549     0.97   0.331    -.0345599     .102461
     lnX1_sq |  -.0029431   .0099295    -0.30   0.767    -.0224046    .0165184
     lnX2_sq |   .0401688   .0146272     2.75   0.006     .0115001    .0688376
     lnX3_sq |   .0437014   .0029943    14.60   0.000     .0378328    .0495701
     lnX4_sq |   .0329704   .0019066    17.29   0.000     .0292336    .0367072
     lnX5_sq |   .0159582   .0021779     7.33   0.000     .0116895    .0202269
     lnX6_sq |   .0177124   .0021018     8.43   0.000      .013593    .0218318
       yr_sq |  -.0082729   .0058903    -1.40   0.160    -.0198178    .0032719
    lnX1lnX2 |  -.0285242   .0094039    -3.03   0.002    -.0469555   -.0100929
    lnX1lnX3 |  -.0143592   .0028112    -5.11   0.000     -.019869   -.0088494
    lnX1lnX4 |   .0053686   .0030461     1.76   0.078    -.0006016    .0113388
    lnX1lnX5 |  -.0094456   .0036895    -2.56   0.010    -.0166768   -.0022144
    lnX1lnX6 |  -.0020347   .0031439    -0.65   0.518    -.0081967    .0041273
      lnX1yr |   .0022486   .0052732     0.43   0.670    -.0080866    .0125838
    lnX2lnX3 |   .0057868   .0035054     1.65   0.099    -.0010836    .0126572
    lnX2lnX4 |  -.0120935   .0036809    -3.29   0.001     -.019308   -.0048791
    lnX2lnX5 |   .0016695   .0046998     0.36   0.722     -.007542     .010881
    lnX2lnX6 |   .0089408   .0037447     2.39   0.017     .0016013    .0162803
      lnX2yr |  -.0091014   .0062509    -1.46   0.145     -.021353    .0031501
    lnX3lnX4 |   .0028722   .0010704     2.68   0.007     .0007743      .00497
    lnX3lnX5 |  -.0010375   .0014344    -0.72   0.469    -.0038487    .0017738
    lnX3lnX6 |   -.003611   .0008933    -4.04   0.000    -.0053619   -.0018602
      lnX3yr |   .0033229   .0030597     1.09   0.277    -.0026741    .0093199
    lnX4lnX5 |  -.0022676   .0009806    -2.31   0.021    -.0041896   -.0003456
    lnX4lnX6 |  -.0009294   .0010955    -0.85   0.396    -.0030766    .0012177
```

lnX4yr	-.0065832	.0021688	-3.04	0.002	-.010834	-.0023323
lnX5lnX6	-.0030654	.0013217	-2.32	0.020	-.0056559	-.0004749
lnX5yr	.002705	.0022614	1.20	0.232	-.0017272	.0071373
lnX6yr	.0100431	.0022635	4.44	0.000	.0056067	.0144795
regi_code1						
2	-.1888439	.0145738	-12.96	0.000	-.2174081	-.1602797
3	-.2238984	.0127395	-17.58	0.000	-.2488673	-.1989294
4	-.2082579	.0177667	-11.72	0.000	-.24308	-.1734358
5	-.1357074	.0144146	-9.41	0.000	-.1639596	-.1074552
6	.030449	.0239983	1.27	0.205	-.0165868	.0774849
_cons	.512928	.2289719	2.24	0.025	.0641514	.9617047
/lnsigma2	-2.314864	.0160385	-144.33	0.000	-2.346299	-2.283429
/ilgtgamma	-100	.	.	.	.	.
/mu	(omitted)					
/eta	(omitted)					
sigma2	.0987796	.0015843			.0957228	.1019341
gamma	3.72e-44	.			.	.
sigma_u2	3.67e-45	.			.	.
sigma_v2	.0987796	.			.	.

Appendix 3.8. Descriptive statistics of some components in tobit model

Variables	Mean	Sd	Minimum	Maximum
Gender (Z_1)	1.127074	0.3330772	1	2
Education (Z_2)	5.365916	3.954689	0	12
Degree (Z_3)	2.3615	0.9677526	1	4
Ethnic (Z_4)	0.4300965	0.4951212	0	1
Irrigation (Z_5)	4.183794	1.520817	1	5
Disaster (Z_6)	2.887717	1.914126	1	5
Land fragmentation (Z_7)	0.4772845	0.2815045	0	0.9338142