EFFECTS OF FARM SIZE AND GREENLEAF MARKETING ARRANGEMENTS ON SMALLHOLDER TEA PRODUCTION EFFICIENCY IN SELECTED COUNTIES IN KENYA

JOSIAH MWANGI ATEKA

K96/CTY/22687/2011

A THESIS SUBMITTED TO THE SCHOOL OF ECONOMICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN ECONOMICS OF KENYATTA UNIVERSITY

JULY, 2018

DECLARATION

This thesis is my original work and has not been presented for a degree or any other award in any other university.

Signature …… **Date…….**

Josiah M Ateka (Mphil, Bsc Agr Econ)

Registration Number K96/CTY/22687/2011

This thesis is an output of a research that was undertaken under our supervision as university supervisors

Signature ……..…….. **Date ……...**

Dr. Perez Onono

Department of Applied Economics

Kenyatta University

Signature ……..…….. **Date ……...**

Prof Martin Etyang

Department of Economic Theory

Kenyatta University

DEDICATION

This study is dedicated to my wife Rebecca and my children Edwin and Sam for their encouragement and support.

ACKNOWLEDGEMENT

This Thesis was written with the support of many people and various institutions. However I will only mention a few because of limitation of space. I first wish to express my deep appreciation to my supervisors; Dr. Perez Onono and Prof Martin Etyang for their guidance, constructive criticisms, support and encouragement. They read numerous drafts and made invaluable comments and suggestions that have helped the evolution of this study from its very beginning to completion. I thank them for their understanding and for always being available whenever I needed their help.

The financial support for the field work for this study was provided by Jomo Kenyatta University of Agriculture and Technology (JKUAT) and the National Research Fund (NRF). I am deeply grateful for their research grants without which the research work would not have been completed. I also wish to appreciate the National Council for Science Technology and Innovation NACOSTI for granting me a research permit to conduct the study. The Agriculture and Food Authority (AFA) and the Kenya Tea Development Agency (KTDA) deserve mention for providing me with secondary data which enriched the study. I also wish to thank the faculty members of the School of Economics at Kenyatta University. The courses provided at the school have made an invaluable contribution to my understanding of economic theory.

Special thanks also go to my fellow class mates who were wonderful colleagues particularly during the course work and preparation for comprehensive exams. I am specifically grateful to Forah Obebo with whom we have walked through the

PhD research journey. The various informal meetings we held to discuss our studies were insightful and beneficial. I also wish to thank my able research assistants Avelline Ondari, Fredrick Were and Julianne Kaveke who helped with the field work and data entry.

To my family, I am greatly indebted to you for standing with me during the entire period of the study. Finally I am greatly indebted to my almighty God for granting me good health to endure the rigor and stress of writing a PhD Thesis in Economics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AFA	Agriculture and Food Authority
ATET	Average Treatment Effects for the Treated
ATMC	Alternative Tea Marketing Channel
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
ESR	Endogenous Switching Regression
EUI	Efficient Unit Isoquant
FFS	Farmer Field School
FIML	Full Information Maximum Likelihood
GDP	Gross Domestic Product
IMR	Inverse Mills Ratio
ITC	International Tea Committee
Kgs	Kilograms
Kgs/ha	Kilograms per hectare
Kgs Pt /ha	Kilograms of processed tea per hectare
KIPPRA	Kenya Institute for Public Policy Research and Analysis
KTDA	Kenya Tea Development Authority
KTDA LTD	Kenya Tea Development Agency
LP	Linear Programming
NNM	Nearest Neighbor Matching

OLS	Ordinary Least Squares
PPS	Production Possibility Set
PSM	Propensity Score Matching
SFM	Stochastic Frontier Model
SLM	Standard Linear Model
TE	Technical Efficiency
TRFK	Tea Research Institute of Kenya
TRI	Tea Research Institute
VIF	Variance Inflation Factor
VRS	Variable Returns to Scale

OPERATIONAL DEFINITION OF TERMS

Farm size	Refers to the geographical land area or holding under tea measured in acres.
Farm	A decision making unit that uses inputs to tea output
Household	A person or group of persons residing in the same compound or homestead (fenced or unfenced) answerable to the same household head and sharing a common source of food and income. Workers residing within the homestead are included as household members.
Productivity	This is the level of output produced by a unit of input. In this study productivity is measured as the physical output divided by land area in acres.
ATMC	The other channels apart from KTDA through which farmers sell their harvested green leaf.
Smallholder	This means a grower who cultivates less than 20 hectares of tea. This is based on the definition provided in the Tea Act.

ABSTRACT

The smallholder tea sub-sector makes an important contribution in the Kenyan economy. Although the subsector has enjoyed relative growth in terms of acreage, output and number of growers, productivity has remained low. The subsector is characterized by huge differentials between actual and potential yields, which imply existence of production inefficiencies. Studies on efficiency in the sector have focused on the regional differences in efficiency, but have not explicitly considered the effects of farm size on efficiency; despite rising concerns from stakeholders about subdivision of tea farms. Moreover, the influences of the alternative tea market channels (ATMCs) which are outcome of the tea market reforms have also not been accounted for in previous studies. This study estimated the level of technical efficiency (TE) and analysed the effects of farm size and marketing arrangements on efficiency in the smallholder tea subsector in Kenya. Using the multistage random sampling approach, data for the study was collected from a cross sectional survey of 525 tea farming households. The level of TE was estimated using the variable returns to scale DEA model, while the analysis of farm size effects was explored using the Fractional Regression (FR) model which accounts for the fractional nature of efficiency scores. The study further applied the Endogenous Switching Regression (ESR) and Propensity Score Matching (PSM) model to investigate the effect of marketing arrangements on TE. The study found that smallholder tea farms were technically inefficient and were operating below their optimal scale. The estimated mean of TE was 0.46 which implies that tea farmers can achieve their current tea output using only 46 percent of their inputs. The study further found that tea farmers in Nyamira County were technically less efficient than their counterparts in Bomet County. The effect of farm size on TE was found to be nonlinear with TE first falling and then rising with increase in farm size. Apart from contributing to literature on the relationship between farm size and efficiency, the study demonstrates that there exists a threshold of farm size (3.93 acres) beyond which increase in farm size leads to an increase TE. The other factors that were found to influence TE were participation in the Farmer Field School (FFS) extension program, the share of family labour applied in tea farming, the age of the farm and the education level of the household head. With regard to tea marketing, it was found that the ATMCs were used by 36.4 percent of the smallholders in the study area. From the ESR and PSM models the study found a consistent result that ATMC participation increases TE in tea production. The study recommends that Agriculture and Food Authority (AFA) and the respective county governments enact regulations that restrict farm subdivisions and implement policies that encourage consolidation of tea farms, support tea replanting to replace aging tea gardens and deepen market reforms in order to increase the farmers' access to ATMC. In addition, efforts should be directed towards promotion of the FFS extension, addressing labour market imperfections and ensuring that the policy formulation process in the tea sector accounts for regional specific heterogeneities that may impact on efficiency.

CHAPTER ONE

INTRODUCTION

1.1 Background

1.1.1 Tea Production in Kenya

The agricultural sector is the backbone of Kenya's economy and a key source of livelihood for majority of the Kenyan rural population (Republic of Kenya, 2010). The sector directly accounts for about 26 per cent of Kenya's Gross Domestic Product (GDP) and another 27 percent indirectly through linkages with manufacturing and other service related sectors (Republic of Kenya, 2015). In addition, the sector contributes about 65 per cent of Kenya's total exports and provides more than 70 percent of the formal and informal employment in the rural areas (KIPPRA, 2013). Analysis of Kenya's economic statistics shows that the agricultural sector has important influence on the performance of the overall economy (figure 1.1).

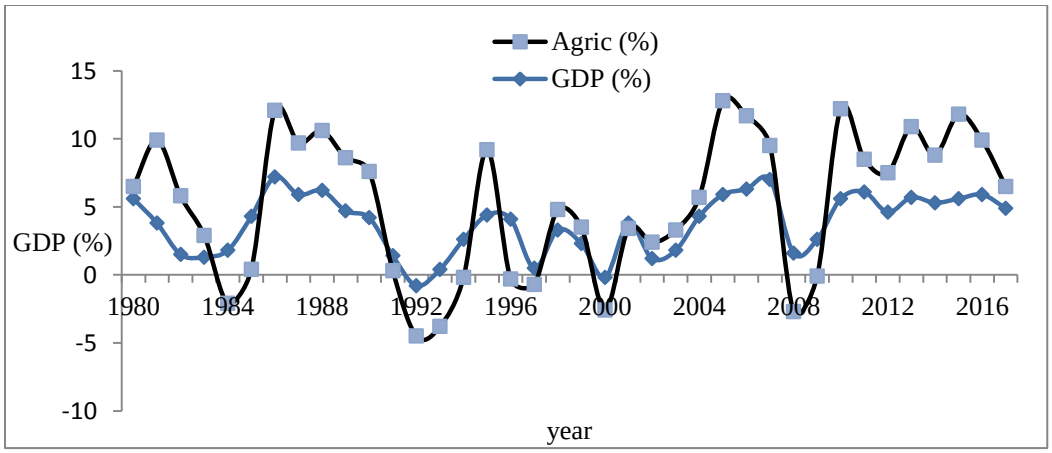


Figure 1.1: Economic and Agricultural Performances in Kenya
Source: Republic Kenya -Economic survey (various issues)

Figure 1.1 shows the trends of economic and agricultural growth in Kenya since 1980. The trends reveal that that the growth of the national economy is positively and highly correlated with growth in agricultural sector. Agriculture is therefore expected to be a key driver in achieving the 10 percent annual economic growth envisaged in the *Kenya's Vision 2030* (Republic of Kenya, 2008; Republic of Kenya, 2010).

Within the agricultural sector, the tea industry makes a key and significant contribution. The industry is a top foreign exchange earner, with the industry earning accounting for about 5 percent and 25 percent of the country's GDP and foreign exchange earning respectively (Republic of Kenya, 2017). The sector directly and indirectly supports over 5 million people, making it one of the leading sources of livelihood in the country (Republic of Kenya, 2015). Globally, the country is the leading exporter of black tea in the world and accounts for 25 percent and 8 percent of world tea exports and production respectively (Onduru *et. al.,* 2012). According to the global industry statistics, Kenya is among the four (4) leading producers; alongside china, India and Sri Lanka, who collectively account for over 75 percent of the global tea production (International Tea Committee, 2013).

Tea growing in Kenya is practised in the country's highlands on the eastern and western sides of the Rift Valley, within altitudes of 1500 to 2700 metres above sea level. These are areas with sufficient and well distributed rainfall and include the regions around Mount Kenya, the Aberdares, and Nyambene hills in central and eastern Kenya and the Kericho, Nandi, Vihiga and Kisii highlands, the Mau

escarpment and Cherangani hills in western Kenya (Kagira *et. al*., 2012; Tea Board of Kenya (TBK), 2014). The leading tea producing counties in Kenya are shown in table A.1 in the appendix.

In Kenya, tea growing has expanded rapidly since its introduction from India in 1903. The country's industry structure is characterized by a dual system comprising of the large scale plantation and the smallholder subsectors (Ogise *et. al.,* 2008). The plantation subsector constitutes the large scale tea estates that are mainly owned by the multinational companies and controls about 40 percent of the industry production (TBK, 2013). The plantation subsector operates alongside a dominant smallholder subsector which produces about 60 percent of the industry production and comprises of over 500,000 smallholder farmers. It is reported to be the biggest and most successful smallholder schemes in the world (Kagira *et. al*., 2012; Onduru *et. al*., 2012). A smallholder within the context of the Kenyan tea sector means a farmer who cultivates tea on less than 20 hectares of land and does not possess his own tea processing factory (Mwaura *et. al*., 2008; Republic of Kenya 2011; Koskei *et. al.,* 2013).

1.1.2 Performance of the Kenyan Smallholder Tea Subsector

The smallholder tea production in Kenya began in the early 1960s following the country's independence. Prior to independence, the colonial policy did not promote smallholder tea production because of the fear that African peasants were not sufficiently skilled; given that tea is a technically very demanding crop and other attempts at smallholder tea production in South Asia had failed (Leonard, 1991). Additionally, discouraging smallholder tea production was consistent with

the need to create a landless class of African peasants to work for the white settlers (Thurston, 1987). Since its inception, the subsector has enjoyed considerable growth in terms of planted area and production as shown in figure 1.2.

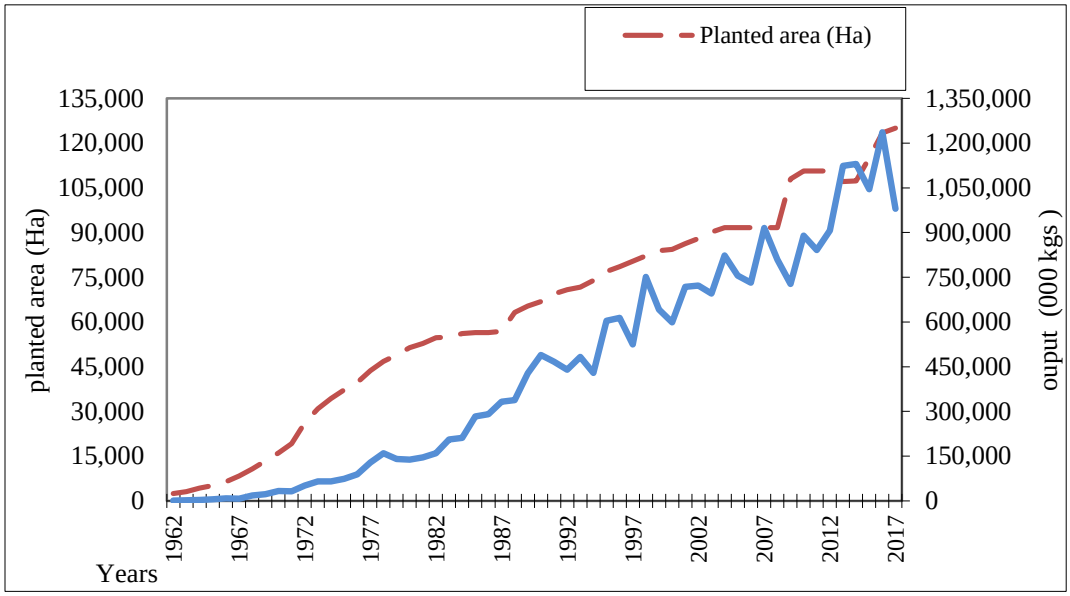


Figure 1.2: Smallholder Acreage and Production Trends (1962-2017)
Source: AFA and KTDA Statistics (1962-2017).

Figure 1.2 shows that the acreage under tea expanded from less than 3000 hectares in 1962 to over 110,000 hectares in 2012; while production rose from about 1.3 million kilograms to over 900 million kgs of green leaf over the same period (Republic of Kenya, 2016). This expansion especially in the earlier years was due to a number of favorable factors including the land distribution policy in the early years of independence, presence of a well-functioning extension system, attractive world market prices and breakthrough in research leading to the release of high yielding clones (Mwaura *et. al.*, 2005; Republic of Kenya, 2016).

Despite its relatively good performance and its immense contribution to the national economy, productivity within the smallholder tea subsector remains low.

Yields in the subsector vary widely marked by huge differentials between the smallholder and the plantation tea subsectors as shown in figure A.1 in appendix 1. The mean annual yields achieved by the smallholder subsector is less than 2000 kgs of processed tea per hectare (kgs pt/ha) compared to the potential of 3525 kgs to 7817 kgs Pt/ha based on the predominant smallholder tea varieties in Kenya (Table A.2 in Appendix 1). Analysis of industry trends, further shows that yields rose impressively in the earlier years after independence but thereafter started to experience setbacks as shown in figure 1.3.

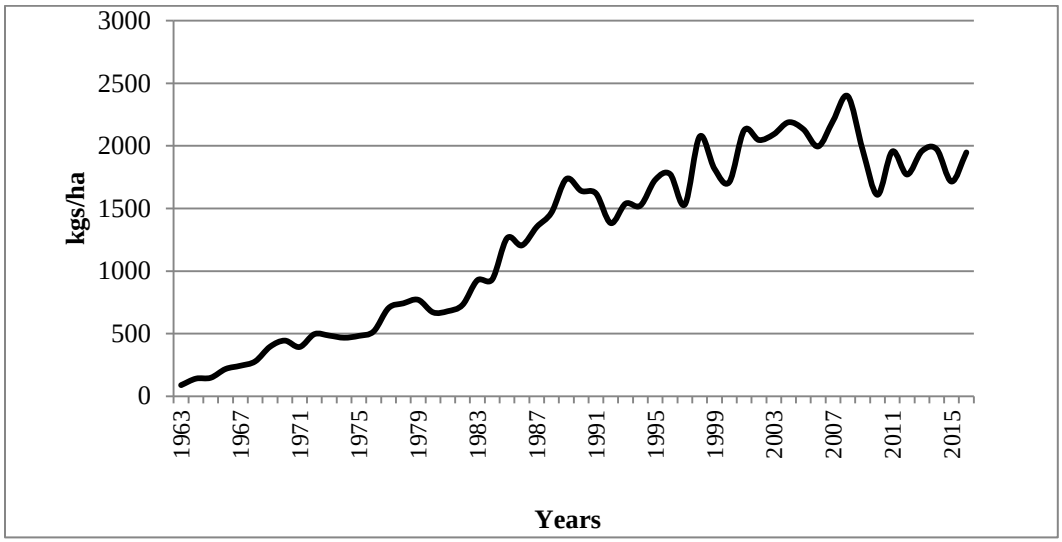


Figure 1.3: Kenya's Smallholder tea Productivity (1963-2017)
Source: KTDA Statistics (various issues 1964-2018)

Figure 1.3 shows that tea productivity in the smallholder subsector was characterized by a steady and consistent increase in yields per hectare from the 1960s to the 1980s. This was however followed by yield fluctuations, interposed by stagnation and decline in the 1990s through to the 2000s. The unimpressive trend coupled with the wide differences between the actual and potential yields point to the existence of inefficiencies and therefore a potential to increase productivity. Enhancing efficiency is an essential factor of productivity growth in

the subsector since tea cultivation requires high investments and involves very high switching costs (Republic of Kenya, 2010).

A part from the tea yields and acreage, the other remarkable trend in the subsector is the sharp rise in the number of tea growers. The trend is attributable to the subdivision of tea farms and largely reflects the pattern of land inheritance and cultural practices on land tenure in the country (Jayne and Muyanga, 2012). In addition to the subdivisions, farms in subsector are also under the threat of conceding a share of their land to other competing enterprises. Following such developments, smallholder tea farms in Kenya are relatively small; with majority of the farmers growing the crop on less than one acre as shown in table 1.1.

Table 1.1: Distribution of Smallholder Farmers by Farm size in Kenya

Farm size (Acres)	No of farmers	Percent (%)	Cumulative (%)
0.0-0.14	45,030.0	8	8
0.14-0.25	95,689.0	17	25
0.25-0.50	253,294.0	45	70
0.50-1.0	118,204.0	21	91
1.0-2.47	45,031.0	8	99
>2.47	5,629.0	1	100
Total	**562,877.0**	**100**	**100**

Source: Kenya Tea Development Agency (KTDA) (2012).

As depicted in the table 1.1, the tea acreage for over 90 percent of the smallholder tea farming households is below one acre. The modal scale ranges between 0.25 to 0.5 acres which is below the estimated minimum of 0.75 acres required to enable a household meet its basic necessities (Ogise *et. al.*, 2008). The scale of farm size is deemed to have significant economic influence on agricultural performance as attested by the many studies that have examined the linkages between farm size and productivity (Balla and Roy, 1988; Barret, 1996; Helfand and Levine, 2004;

Kiani, 2008). Within the context of land constraints in Kenya's tea sector, questions regarding how farms of different sizes differ in their relative efficiency are therefore becoming increasingly relevant (Republic of Kenya, 2014a).

1.1.3 Reforms in the Kenyan Tea Subsector

Tea was first introduced in Kenya on experimental basis in 1903, but commercial tea cultivation began in 1924 (Owuor, 1999). In order to govern it's production, the first legal instrument to be enacted was the Tea Ordinance of 1934 (Republic of Kenya, 2007). The ordinance was revised in 1948 creating the institution of the Directorate of Agriculture which was responsible for controlling the production of tea by issuing permits to growers. This continued until 1950 when the Tea Board of Kenya (TBK) was established under the Tea Act (CAP. 343) to regulate the industry. Through the Act, the Kenya Tea Development Authority (KTDA); a state corporation established in 1964 was given exclusive control in the provision of extension, leaf collection and marketing services. This organizational design reflected the belief that the services required for tea production would not be supplied through market forces (Leonard, 1991). The system of controls continued until 1999, when the tea industry was liberalized (Kinyili, 2003).

As part of the liberalization, the key change that was instituted was the privatization of KTDA. This led to the deregulation of markets following protests by farmers that the fortunes of the tea sector had started to diminish after impressive development in the first two decades after inception (Nyangito and Kimura, 1999; Mbeche, 2012). The main complaints were that the KTDA system lacked accountability and was inefficient, as evidenced by the poorly coordinated

tea collection and processing operations (Nyangito, 2000). Despite attempts by the government to protect KTDA as a strategic parastatal, privatization was completed in 2000 culminating in the restructuring of KTDA into a private farmers' company (Mbeche, 2012). The other changes that took place include the repeal of the tea growing license and cess and the removal of subvention of staff seconded to the tea factories (Republic of Kenya, 1999). These reforms were implemented to promote efficiency and competitiveness through competition and involvement of farmers in management of the subsector (Republic of Kenya, 2007).

A part from liberalization, the other major regulatory amendments implemented in the industry include the Tea Trade Regulations of 2008 and the Tea Amendment Bill of 2011. The amendments were introduced to align the existing legal instruments with liberalization in the sector and accommodate the recommendations of the Tea Industry Task Force Report of 2007. Key changes included the Repeal of prohibition to uproot tea, introduction of green leaf supply agreements and introduction of guidelines and clarification of roles of managing agents (Republic of Kenya, 2011).

Despite its critical role in the national economic growth, the smallholder tea subsector has continued to experience challenges such as high cost of inputs, technology transfer bottlenecks, lack of a proper policy and regulatory framework, over reliance on few traditional tea markets, fluctuation in producer prices and lack of accountability and poor information sharing among stakeholders (Gesimba *et. al.,* 2005; Republic of Kenya, 2007; Kagira *et. al.,* 2012; Njeri *et. al.,* 2013). The presence of challenges in the subsector is evidenced by the low productivity

and rising concerns that reforms in the subsector have not significantly improved the livelihood of tea farmers (Republic of Kenya, 2007; Mbeche, 2012).

1.1.4 Green Leaf Marketing Systems in Kenya

Following liberalization of the smallholder tea subsector in Kenya, considerable changes have taken place in the industry. One key development is the emergence of alternative tea market channels (ATMC) for selling green leaf, alongside the more established KTDA system. Unlike in the KTDA system where green leaf from the farms is sold at designated buying centres, transactions in the ATMC mainly involve intermediaries at the farm-gate or roadside (Nyangito, 2000). The alternative system is often characterized by immediate cash payments and mainly operates in Kiambu, Kericho, Bomet, Kisii, Nyamira, Nandi, Kakamega and Vihiga counties (Kinyili, 2003). In contrast, payments under the KTDA system are structured into an initial monthly and end of year payment also known as bonus.

Since the ATMC is an outcome of reforms which were introduced to promote efficiency, an empirical understanding of its effects on efficiency is justified. This is in light of apprehensions that reforms in the Kenyan tea sector have not improved the farmers' fortunes (Sanne van der Wal, 2008; Mbeche, 2012). The understanding is also important in view of findings from other countries showing that institutional arrangements have significant effects on productivity in tea production (Chirwa and Kydd, 2006; Yoshiko, 2011). While the performance implications of the ATMC have been debated by stakeholders including the National Assembly (Omiti and Mosoti, 2010), empirical evidence linking the alternative system with efficiency is needed to appropriately inform policy.

1.2 Statement of Problem

Although the performance of the smallholder subsector has been characterized by substantial growth in acreage, production and number of growers, productivity has remained low (Ogise *et. al.,* 2008; Kagira *et. al.,* 2012). Productivity trends in the subsector indicate a phenomenal growth from inception in early 1960s up to the 1980s. Thereafter, productivity started to diminish in the early 1990s marked by a huge gap between actual yields and the industry potential. Despite the implementation of reforms to reverse the trend, the situation has prevailed to date. The huge differential between the actual and potential yields imply the existence of inefficiencies and opportunities to improve the welfare of the smallholder tea farmers. The smallholder tea subsector in Kenya is characterized by small farm holdings following rapid sub-division of tea farms. This has led to rising concerns that the decline in farm units would be a cause of stagnation in tea yields. Envisaged reforms to address the concerns include the consolidation of farms and enterprise diversification to support farmers with very small holding to exit the industry (Republic of Kenya, 2014a). Understanding the efficiency behaviour among smallholders and its determinants as well as the importance of the effect of farm size is critical for the formulation and implementation of required reforms.

Studies that have analysed efficiency in the tea sector in Kenya have focused on comparison of regional differences in efficiency between farmers in the East and the West of the Rift Valley (Kavoi *et al,* 2001; Ogise *et. al.,* 2008; Kiprono, 2013). The studies have provided insights on some of the determinants and the reasons for the regional disparities in efficiency, but have not explicitly considered the

effects of farm size on efficiency. While the investigation of the relationship between farm size and productivity is a common empirical regularity (Dyer, 1997; Helfand and Levine, 2004; Rios and Shively, 2005; Mugera and Langemeir, 2011; Ali and Deininger, 2014; Heath, 2015), the subject has received less attention in Kenya. Consequently, despite rising concerns about subdivision of tea farms, the economic rationale for policy intervention or restrictions has not been established.

The advent of the alternative green leaf markets is one of the key developments in the industry following liberalization of the sector. The alternative markets operate in some parts of the tea growing areas in the country alongside the dominant KTDA system. Although it was envisaged that liberalization would trigger competition and spur efficiency, the influence of the alternative markets on efficiency has not been accounted for in previous studies. Studies that have assessed the impact of liberalization of the tea industry in Kenya such as Nyangito (2001) and Mbeche, (2012) have mainly focused on accountability, governance and value chain challenges in the subsector. In addition, other studies have analysed the determinants of participation in the alternative markets (Chepkulei, 2013). As a result the industry lacks an empirical basis for enacting appropriate institutional arrangements to guide the marketing of green leaf among the smallholder farmers so as to address the problem of absence of a suitable regulatory framework for green leaf sales as has been identified in various tea industry reports (Republic of Kenya, 2007; Republic of Kenya, 2014a; Republic of Kenya, 2016).

This study therefore analysed efficiency in Kenya's smallholder tea subsector by incorporating the effects of farm size and green leaf marketing in the models that were estimated. The study also ascertained the influence of other variables on efficiency.

1.3 Research Questions

The following questions have been addressed in this study;

i) What is the level of technical efficiency in smallholder tea production in selected counties in Kenya?

ii) What is the effect of farm-size on technical efficiency in smallholder tea production in selected counties in Kenya?

iii) What is the effect of green leaf tea marketing arrangements on technical efficiency in smallholder tea production in selected counties in Kenya?

1.4 Objectives of the Study

The general objective of this study was to analyse the effects of farm-size and green leaf marketing on the technical efficiency of smallholder tea production in selected counties in Kenya. The specific objectives of the study were to:

i) Estimate the level of technical efficiency in smallholder tea production in selected counties in Kenya.

ii) Determine the effect of farm-size on technical efficiency in smallholder tea production in selected counties in Kenya.

iii) Analyse the effect of green leaf tea marketing arrangements on technical efficiency in smallholder tea production in selected counties in Kenya.

1.5 Significance of the Study

This study focused on the analysis of technical efficiency in the smallholder tea subsector in Kenya. The research is important based on the following perspectives.

First, given the importance of smallholder tea subsector in the Kenyan economy, the estimation of technical efficiency reflects the existing potential to improve productivity within the existing resource levels and provides answers to questions on what factors hold back smallholders from increasing their productivity. An understanding of the relationships between efficiency and the household and farm-specific variables provides KTDA and other industry players with information to design programs that can help increase productivity among farmers.

Secondly, within a context of rising concerns about subdivision of tea farms, the findings of the study provide an empirical basis for policy reforms on farm size in the Kenyan smallholder tea subsector. In addition, the study is important to academicians since it contributes to the body of knowledge on the unsettled debate on the relationship between farm size and efficiency. The understanding of how farms of different sizes differ in their relative efficiencies is crucial to policy questions on the future farm-size structure, supply response and international competitiveness (Gorton and Davidova, 2001).

Finally, the research has provided the industry regulator with insights on suitable arrangements for green leaf marketing in the smallholder tea farmers.

1.6 Scope and Limitation of the Study

This study sought to examine the effects of farm size and green leaf marketing arrangements on efficiency in the smallholder tea subsector in Kenya. The study was conducted in two counties in Kenya, namely Bomet and Nyamira. The counties are among the country's leading producers of tea and were purposely selected since they represent areas where the alternative tea marketing system which was under investigation in the study thrives. In addition, the counties provide a fair reflection of the variability in farm size across the tea growing areas in Kenya (Table A.1). The scale of holding in the two counties has representation of both the small and relatively large tea farms. In addition, the selected counties have relatively similar agro-ecological conditions, which was necessary in order to minimize the effects of geo-climatic variability in the analysis of efficiency.

Although this study was sought to analyse the effects of farm-size and green leaf marketing on the technical efficiency of smallholder tea production in Kenya, it was not possible to carry out the study in all the tea growing counties in the country. Undertaking such a study would require massive financial and time resources. As such, the study was limited to two counties, implying that some aspects of study's findings may not be fully generalized to the other counties.

1.7 Organization of the Thesis

This thesis is organized into five chapters. Chapter one presents an introduction of the background to the problem and the need for the study. The chapter also presents the statement of the problem, the study objectives, research questions, significance and scope of the study. Chapter two of the thesis contains a review of

literature and includes a discussion of the theoretical issues and empirical studies that are relevant to the study. The chapter also includes a consideration of the models applied in the measurement of efficiency and estimation of treatment effects and concludes with an overview of the review of literature.

The third chapter of the Thesis describes the research design and the theoretical and empirical models. In addition, the chapter presents a description of the definition and measurement of variables, the sample size and sampling procedures and methods of data analysis. The results and findings of the study are discussed in chapter four while chapter five concludes the thesis with a summary of findings, conclusions, policy implications, contribution to knowledge and areas of further research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter provides a review of existing theoretical and empirical literature that is relevant to this study. The chapter has three sections. The first section presents the theoretical literature in whose context the study is set. The second section presents a review of empirical studies and the last section concludes the chapter with an overview of literature.

2.2 Theoretical Literature

This section provides a discussion of the theoretical literature and is organized as follows. The first part reviews the agricultural production theory for purposes of understanding the characterization of agricultural production technology and therefore the basis for estimating efficiency. The second part discusses the approaches and models for estimating technical efficiency (TE). The third part reviews the theory of utility maximizing peasant for purposes of understanding the household demand for green leaf market channels. The last part reviews the models for estimation of treatment effects which provided the basis for modeling of the efficiency effect associated with the participation in the alternative tea market channels (ATMCs) for selling green leaf.

2.2.1 The Agricultural Production Theory

2.2.1.1 Specification of Farm Technology

The agricultural production theory describes agricultural production as a process that utilizes technology to transform a set of inputs into output (Heady and Dillon, 1969; Just and Pope, 2001; Debertin, 2002). The state of technology defines and restricts what is technologically feasible in combining inputs to produce output and can be characterised by a production possibility set (PPS). Following Lovell (1993), the PPS of a farm using a vector of N inputs to produce vector of M outputs can be represented as;

$$T = \left((X, Y) : X \text{ can produce } Y\right) \in \Re^{M+N} \qquad (2.1)$$

Equation 2.1 (PPS) describes the collection of all the vectors of output (Y) that are feasibly producible using the input vector X. The PPS describes the technologically feasible production plans of a farm in a more general way, since it allows the description of multiple input and output technologies (Varian, 1992; Dinh, 2011). In the case of a farm producing a single output from multiple inputs; it is more convenient and common to describe the farm's technology in terms of a production function (Jehle and Reny, 2011), which can be specified as;

$$y = f(X) \qquad (2.2)$$

Where, y is the maximum quantity of output producible from a set of inputs *X*, given the available technology (Heady and Dillon, 1969; Varian, 1992). In

agricultural production the vector X will typically include labour, land and other purchased inputs such fertilizer and seed.

The properties of equation 2.2 are specified by the assumptions that X is a set of real non-negative inputs and $f(X)$ is finite, non-negative, continuous, strictly increasing and strictly quasi-concave (Chambers, 1988). Continuity of the function ensures that small changes in inputs results in small changes in the quantities of output produced while the assumption that f(X) strictly increases in the input vector (X) ensures that using more of each input results in strictly more output. The strict quasi-concavity of the function implies presence of input complementarity in production such that in a situation where two inputs are applied in production, the average of two extreme input vectors will produce strictly more output than at least one of the two extreme input vectors (Jehle and Reny, 2011).

The behavioral relationship between inputs and output can further be characterized by returns to scale (RTS) in production (Dinh, 2011). The farm's technology can exhibit constant returns to scale (CRS) or variable return to scale (VRS). The CRS production technology describes a situation where a given percentage increase in inputs leads to the same percentage rise in output. The VRS production technology applies where a given percentage increase in inputs leads to a less or more than proportionate increase in output hence yielding an increasing or decreasing VRS. In the increasing VRS, output rises by a higher percentage than the percentage rise in inputs; while in the decreasing VRS the percentage rise in output is smaller than

percentage rise in inputs (Daraio and Simar, 2007). The representation of returns to scale is important in agricultural production analysis since it indicates whether any efficiency gains can be obtained by adjusting the scale of operation of a farm (Tolga *et. al.*, 2009).

A part from the properties that define the shape of the production function, a satisfactory specification of agricultural production technology should provide specific attention to the unique features that differentiate agriculture from the other sectors (Debertin, 2002; Karagiannis, 2014). Agricultural technology is associated with its biological nature and exposure to widely varying elements of the biophysical environment which often are beyond the control of the smallholder farmers (Just and Pope, 2001). The set of variables related to the biophysical environment such as soil quality and other agro-climatic conditions can exert significant influence on production and are hence included in agricultural technology specifications (Karagiannis, 2014). Following Just and Pope, (2001), a general agricultural production function is therefore specified as;

$$Y = F\left(K, X^{ti}, Z^{ti}, \omega \middle| \sum_{i} Z^{ti} \leq \bar{M} \right) \tag{2.3}$$

Where, Y is a vector of agricultural outputs, K is vector of fixed inputs such as land, buildings and machinery, ω is vector of the biophysical variables capturing the influence of geo-climatic heterogeneities, and X^{ti} and Z^{ti} are vectors representing purchased and farmer controlled variable inputs respectively. The superscript *ti* reflects the seasonality and sequential nature of agricultural

production stages imposed by biological characteristics. Such a recursively separable structure of the agricultural production process implies for instance that labour applied for pre-harvest activities such as planting and weeding is separable from labour applied to harvesting activities. The vector $\bar{M}$ represents the maximum use or availability of services made possible by the fixed stock of farmer-controlled inputs in each stage of the production process. The equation emphasizes the fact that there are a number of unique features that typify agricultural production and while some parallels can be drawn with other sectors, the extent in which the features occur in agriculture has implications on how they can be represented empirically.

2.2.1.2 Efficiency in Farm Production

Theoretically, a production function represents the limit or boundary of the PPS such that a productive farm operating on its production function can be considered to be efficient in the use of its inputs (Pascoe *et. al.,* 2001). Efficiency in this context therefore measures the performance of the farm in terms of its ability to operate close or on the boundary of its PPS or frontier (Green 2003). This means for instance that if from a given set of inputs there exists a maximum level of output that is producible, the level of efficiency of a farm could be measured by comparing its level of output to that of a farm realizing the highest output using the same set of inputs. From a theoretical perspective, the proposition of efficiency ascribes to the household a behavioral motivation of profit maximization and an economic-technical component in terms of the farm's performance as an enterprise (Schultz, 1964; Mendola, 2005).

Since efficiency is a broad concept, its empirical investigation has led to the need to differentiate between two concepts of economic efficiency; Allocative efficiency (AE) and technical efficiency (TE) (Yotopolous and Lau, 1971; Coelli *et. al.,* 2005). TE refers to the ability of the firm to obtain maximum output with the available resources given the range of alternative technologies available to the farmer (Rios and Shively, 2005; Mugera and Langemeier, 2011). In contrast, AE refers to the ability of a firm to make adjustment on inputs and outputs to reflect relative prices. The latter requires the attainment of the marginal conditions for profit maximization. Therefore the estimation of TE only requires input and output data while AE requires price data as well (Vincova, 2005). This study was concerned more with productivity based on physical input-output relationships; hence the relevant concept was that of TE, which is considered to be a good measure of total factor productivity for a single output- multiple input production situation such as in tea production (Green, 2003).

According to Debreu (1951) and Farrell (1957) TE can be analysed using deviations from an idealized isoquant. The argument is that although a production function theoretically represents the boundary of the PPS, an empirically estimated production function only shows the average technology and not a frontier against which efficiency can be analysed (Green, 2003). Following Farrell (1957), TE can be conceptualized as the extent to which a farm's actual inputs can be contracted towards the boundary of the PPS (the idealized isoquant), while holding the level of output constant as shown in figure 2.1.

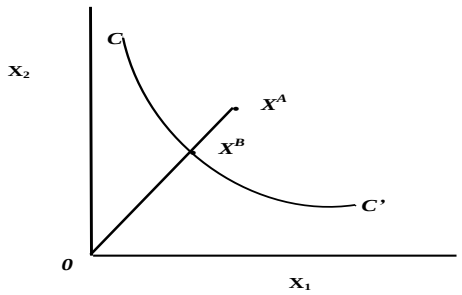


Figure 2.1: Conceptualization of TE Based on an Idealized Isoquant

Figure 2.1, shows an isoquant in which the production technology of the firm is assumed to be characterized by existence of a smooth, continuously differentiable, constant returns to scale (CRS) and a strictly quasi concave production function (Jehle and Reny, 2011). The input vectors X^A and X^B represent the combination of inputs used by two farms A and B to produce a unit of output. Assuming that the curve CC' is the efficient unit isoquant (EUI), then X^B represents an efficient input set for producing a unit of output while X^A is an inefficient input set. The TE of farm A would be represented by the fraction $\frac{OX^B}{0X^A}$ since it represents the proportion of inputs that an efficient farm (in this case farm B) uses to produce the same level of output (Farrell, 1957; Coelli *et. al.,* 2005). TE in this case shows the possible proportional reduction of inputs that can be achieved for farm A without any reduction of its output. The input oriented measure of TE therefore assumes a value of one for an efficient firm and becomes small with increase in the amount of inputs a farm uses to produce a unit of output.

The implication of Farell's model is that the empirical estimation of TE would begin with the identification of an efficient isoquant (EI). This is because the

boundary of the PPS which production theory associates with any production activity is not known in practice and can only be derived from observations of inputs and output from farms in an industry (Coelli, 1996). Using the identified EI, TE is then obtained by comparing the input set of a farm operating on the EI with the input set of a farm whose TE is being evaluated. The approaches commonly applied for the identification of the EI and hence the empirical estimation of TE are discussed in the subsequent section.

2.2.2 Approaches for Estimation of Technical Efficiency

Two broad approaches namely, the Parametric Stochastic Frontier Model (SFM) and the non-parametric Data Envelopment Analysis (DEA) are commonly applied in the empirical estimation of TE.

2.2.2.1 The Stochastic Frontier Model (SFM)

The SFM is an econometric approach for estimating TE based on the analysis of regression residuals (Green, 2003). The model is founded on the theoretical premise that a production function represents the maximum output attainable given a set of inputs, and the intuition that the production process is subject to a composite error term, comprising of two independent and distinguishable disturbances (Aigner, *et. al.,* 1977; Meeusen and van den Broeck, 1977). One of the disturbances captures the stochastic effects that are outside the farmer's control such as the weather and natural shocks, measurement errors and other statistical noise and the other accounts for the presence of technical inefficiency (Obare *et. al.,* 2010).

SFM therefore involves estimation of a production equation where a farm's output is a function of inputs and the disturbance which can then be decomposed into its components as shown in equation 2.4.

$$Y_i = F(X_i;\beta) + \varepsilon_i \text{ such that } \varepsilon_i = v_i - u_i \qquad (2.4)$$

$$i = (1, 2, \ldots, n)$$

Where Y_i is a vector of output, X_i is a vector of inputs, β is vector of parameters and ε_i is the residual term defined by two terms (v_i and u_i) which are assumed to be identically and independently distributed. In the model (equation 2.4), v_i is a two-sided normally distributed random error [$v \sim N(0\ \sigma_v^2)$] while the technical inefficiency term u_i is assumed to be one sided non-negative term $\left[u \sim N^{+}(0\ \sigma_u^2)\right]$ with a half-normal distribution (Green, 2003). The non-negativity condition ($u_i \geq 0$) implies that all the observations of Y_i are constrained to lie on or beneath stochastic frontier, such that the i[th] farm is said to be technically efficient if $u_i = 0$ and inefficient if $u_i > 0$ (Battese and Coelli, 1988).

To obtain estimates of TE using this approach, it is necessary to decompose the inefficiency component from the compounded error term. Following Jondrow *et al.* (1982), the decomposition is achieved by taking the conditional expectations of u_i given ε_i as shown in equation 2.5.

$$E(u_i|\varepsilon_i) = \sigma_* \left\{ \frac{f\left(\frac{\varepsilon_i \lambda}{\sigma}\right)}{1 - F\left(\frac{\varepsilon_i \lambda}{\sigma}\right)} - \left(\frac{\varepsilon_i}{\sigma}\right) \right\} \qquad 2.5$$

Where $\lambda = \left(\frac{\sigma_u}{\sigma_v} \right)$ and f (.) and F (.) are the normal probability density and distribution functions, respectively. According to Battesse and Corra (1977), consistent estimates of λ, σ and β can be obtained from the maximum likelihood estimation of equation 2.4. Consequently, the conditional mean of the distribution u_i (equation 2.5) can be used as point estimate for technical inefficiency.

Although the SFM approach has the capability to explicitly separate inefficiency from stochastic noise (Karagiannis, 2014), the methodology is prone to misspecification due to the imposition of explicit assumptions on the functional form and distribution (Färe *et. al.*, 1994). Further, the use of SFM was discounted in the present study on the basis that it is not suitable for the two step approach which had to be adopted in this study to address the research objectives. Doing so would involve contradictory assumptions about the distribution of the inefficiency term.

2.2.2.2 The Non-Parametric Approaches: Data Envelopment Analysis (DEA)

The Data Envelopment Analysis (DEA) is non-parametric linear programming (LP) specification that involves comparison of observed producers with each other. Like SFM, the model is premised on the existence of a production frontier or a best practice technology and variations in performance in the transformation of inputs and outputs among producers in an industry (Green, 2003). DEA involves fitting a linear quasi convex hull around the input –output data of observed farms then determining TE as each farm's distance from it (Daraio and

Simar, 2007). The model is deterministic and therefore assumes that any deviations from the hull are attributed to inefficiency.

DEA as a model was operationalized as a LP estimator by Charnes *et. al.* (1978) and its essence in measuring the efficiency of a farm lies in maximizing its efficiency rate. The model assumes that each farm attaches its own weights (relative importance) to both inputs and outputs, such that the efficiency rate of a farm can be expressed as the ratio of weighted sum of outputs and inputs (Ogada *et. al.*, 2014). Therefore, with N decision making units, DMU_1 DMU_2,…, DMU_n, where each unit produces a vector of *M* outputs using *K* number of inputs; the efficiency level of the Q^{th} *DMU* can be computed as ;

$$\text{Efficieny rate (ER)} = \frac{\sum_{i=1}^{m} U_i Y_i^Q}{\sum_{j=1}^{K} V_j X_j^Q} \qquad (2.6)$$

Where; V_j (j = 1, 2, …, K) represents the weight assigned to j^{th} input, U_i (i = 1, 2, …, M) is the weight assigned to i^{th} output and X_j are vectors of inputs used in the production of output (Y_i) by the Q^{th} farm. Since, efficiency of a given DMU is assessed relative to the other farms; the optimal weights (V_j's and U_i's) can be derived by solving the following problem;

$$\text{Maximize (ER)} = \frac{\sum_{i=1}^{m} U_i Y_i^Q}{\sum_{j=1}^{K} V_j X_j^Q} \qquad (2.7)$$

Subject to $$\frac{\sum_{i=1}^{m} U_i Y_i^Z}{\sum_{j=1}^{K} V_j X_j^Z} \leq 1 \quad (2.8)$$

$$U, V \geq 0 \text{ and } Z = 1, 2, ..., n \quad (2.9)$$

The solution of the problem in equation 2.7 to 2.9 solves for values of *U* and *V* that maximize the efficiency of the Q^{th} farm, subject to the condition that efficiency rate of any other DMU (Z) in the sample must not be greater than one (Vincova, 2005). The efficiency rate for the Q^{th} farm can be obtained by substituting the optimal values of U and V into equation 2.6. However a problem with the resultant ratio is that it is not unique since infinite set of solutions are feasible (Coelli, 1996). The problem is avoided by imposing the condition $V_j X_j = 1$ and transforming the model into a LP specification in matrix notation as shown in 2.10 to 2.13.

$$\text{Max } U'Y \quad (2.10)$$

Subject to;

$$V'X = 1 \quad (2.11)$$

$$U'Y - V'X \leq 0 \quad (2.12)$$

$$U, V \geq 0 \quad (2.13)$$

Where U' is a M x 1 vector of output weights, Y is M x N output matrix V' is a M x 1 vector of input weights and X is a M x N input matrix. The specification in equation 2.10 to 2.13 is known as the primary or the multiplier form of the LP model and is commonly referred to as CCR LP after Charnes, Cooper and Rhodes

(1978). In most empirical applications, the stated LP specification has been solved using its dual which is more preferable since it has fewer (K+M) constraints than its primal which has N+ 1 constraints (Daraio and Simar, 2007). The dual specification following (Coelli (1996), which was also applied in this study is shown in equations 2.14 to 2.17.

$$\text{Minimize}\,\theta \qquad (2.14)$$

Subject to;

$$-Y^{q} + \lambda Y \geq 0 \qquad (2.15)$$

$$\theta X^{q} - \lambda X \geq 0 \qquad (2.16)$$

$$\lambda \geq 0 \qquad (2.17)$$

The model solves for values of θ; where θ represents the proportion of input bundle of the q^{th} farm needed to produce its own output by the virtual(frontier) farm and λ is $(\lambda = \lambda_1, \ldots, \lambda_n)$ is a Nx1 vector of constants that shows the intensity with which each farm is used in the construction of the frontier farm. The model is interpreted as seeking a virtual farm characterised by inputs λX and outputs λY, which are a linear combination of X and Y of the other farms in the industry which are better than the farm being evaluated; such that $\lambda X_j^{virtual} \leq X_j^{Q}$ and $\lambda Y_i^{virtual} \geq Y_i^{Q}$ (Vincova, 2005). The Q^{th} farm is therefore considered efficient if its output is best produced using all of its inputs, or inefficient if its output can be produced by the virtual farm, using a fraction θ of its inputs.

The model is known as an input oriented DEA model since it seeks to improve the input characteristics of the farms under evaluation (Coelli 1996). Its alternative,

the output oriented DEA model assesses whether a given DMU can increase its output without increasing the input vector. Although the decision on which orientation to use depends on which factors (inputs or outputs) the farmer has more control over, there are some studies that have shown that the adopted orientation only adds minor influences on the estimated efficiency (Coelli and Perelman, 1996).

The shape of the frontier obtained from DEA can be viewed as exhibiting constant CRS or VRS technologies. Returns to scale therefore provides another common approach for categorizing of DEA models. The most common models are the CRS model proposed by Charnes *et. al.* (1978) and the VRS model proposed by Banker *et. al.* (1984). The CRS model suffers from the weakness that it assumes that all farms operate on the optimal scale, which is not the case in reality (Dinh, 2011). The VRS model on the other hand is flexible and makes it possible to explore scale issues within DEA. This is achieved by adding a convexity constraint ($\sum \lambda = 1$) into the standard CRS specification (equation 2.14 to 2.17) in order to ensure that the farms under evaluation are only compared against farms of similar size (Ji and Lee, 2010; Muniu *et. al.*, 2015).

Although the DEA approach suffers from an inability to explicitly separate inefficiency from stochastic noise (Asefa, 2011), it was considered in the present study since it does not impose parametric restrictions on the underlying technology and is therefore less prone to misspecification (Färe *et. al.*, 1994; Zhu, 2014). The use of DEA was also motivated by the fact that smallholders are very heterogeneous making uniformity in their production function unlikely and

therefore rendered SFM less suitable (Ogada *et. al.,* 2014). In addition, DEA provides other useful information on returns to scale and optimal scale of productive units (Ji and Lee, 2010). Finally in the case of tea production, DEA has the additional advantage that it can facilitate the identification of efficient tea farming practices by making reference to the peer groups of best performing farms (Dinh, 2011). This study employed the VRS DEA which allows for investigation of scale properties by decomposing TE into two mutually exclusive and non-additive components; Pure TE and scale efficiency (SE). Pure TE reflects the managerial performance of the farm in organizing inputs in the production process, while the latter component reflects the farm's ability to choose the optimal size (Kumar and Gulati, 2008).

2.2.3 The Theory of the Utility Maximizing Peasant

The theory of the utility maximizing peasant was developed by Chayanov in the 1920s. The theory was developed to capture the behavior of agricultural households and takes account of the dual character of the peasant household as both family and enterprise and therefore the production and consumption side of the model (Chayanov, 1966). In the model, households are assumed to make production, resource allocation and consumption decisions to maximize utility subject to a set of constraints as shown in equation 2.18 to 2.21 following Singh *et. al.*(1986) and Ellis (1992).

$$\text{Max}\, U(C_F, Q_{NF}, \ell, Z^h) \quad (2.18)$$

Subject to;

$$F(Y_F, Q_C, L, X) = 0 \quad (2.19) \text{ (Production)}$$

$$T = H + L + \ell \quad (2.20) \text{ (Time endowment)}$$

$$P_F(Y_F - C_F) + W(H) + P_C Q_C = P_X X + P_{NF} Q_{NF} \quad (2.21) \text{ (Cash income)}$$

Where; equation 2.19, 2.20 and 2.21 are the production, time endowment, and cash income constraints respectively; C_F, Q_{NF} and ℓ represent the consumption of food, non-food commodities and leisure respectively, while Z^h represents household characteristics influencing the marginal utilities of consumption. Y_F is the output of food, while Q_C is cash crop (tea) output; L is the household labor allocated to the production of the food and cash crop and X is vector representing the other purchased agricultural inputs including hired labour. The household endowment of time in each season is represented by T and is allocated between leisure (ℓ), household labour (L) and hours of work for pay (H) according to Becker (1965) and Taylor and Adelman (2003). W is the wage rate of labour and P_F, P_C, P_X and P_{NF} are the market prices for food, cash crop (tea), other agricultural inputs and the nonfood items respectively.

The constraints in the maximization problem can be combined into a single "full income" constraint (Ellis, 1992; Huffman, 2010), as shown in equation 2.22.

$$\left((P_F Y_F) + (P_C Q_C)\right) - (P_X X + WL)\big) + WT = \left(P_F C_F + P_{NF} Q_{NF} + W\ell\right) \quad (2.22)$$

The full income constraint shows that the full income for the household income is received from farm profits (the first term on the left side of equation 2.22) and the sale (value) of the household time endowment (T) at the wage rate (W). The right side of equation 2.22 shows that the full household income is spent on leisure $W\ell$ and the purchase of food ($P_F C_F$) and non-food items ($P_{NF} Q_{NF}$). The problem of the household is therefore to maximize its utility subject to the full income constraint. Using the Lagrange multiplier (λ) approach, the first order necessary conditions (FONC) for the augmented utility maximization problem yields a system of equations specified in equations 2.23 to 2.26;

$$U'(C_F) = \lambda P_F \quad (2.23)$$

$$U'(C_{NF}) = \lambda P_{NF} \quad (2.24)$$

$$U'(\ell) = \lambda W \quad (2.25)$$

$$P_F Y_F + P_C Q_C - (P_X X + WL) + WT - P_F C_F - P_{NF} Q_{NF} - W\ell = 0 \quad (2.26)$$

Assuming that interior solutions exist, equations 2.23 to 2.26 can be solved jointly to obtain a general form of the household demand functions for leisure (ℓ), food (C_F) and non-food items (Q_{NF}) (de Janvry *et. al.,* 1991). These demands are functions of the price of food (P_F), price of non-food items (P_{NF}), wage rate (W), and farm profits which among other variables depends on the price and level of inputs (X) and the price and output of the cash crop (Q_C) (Huffman, 2010).

An extension of the utility maximization model allows for analysis of market participation decisions by farm households. An example is Key *et. al.* (2000) who showed that in the presence of high transaction costs on food markets, households

have to endogenously choose market participation strategies where they may enter markets as seller or buyers, or remain in autarky, only producing what they consume. The extension entailed modification of the full income constraint to reflect the fact that transactions costs lower the price effectively received by sellers and raises the effective price paid by buyers. Consequently, a part from the optimal input and output demands, the solution of the utility problem also contained optimal choices on market participation.

Following Fafchamps and Hill, (2005) and Mmbando (2014), the extension of the model to accommodate transaction costs in the case of a cash crop such as tea would need modification of the second term of equation 2.22 $(P_C Q_C)$ as follows;

$$\left(P_C^{Di} - TC^{Di}\right)Q_C \qquad (2.27)$$

Where P_C and Q_C (as defined earlier) are respectively, the price and quantity of the cash crop, and TC represents the transaction costs faced in the cash crop market. The superscript $(D\text{i})$ is a market participation indicator reflecting the choice of a market channel through which the household sells its produce. Equation 2.27 shows that the effective cash crop price received by the household depends on choice of a market channel, the price offered in the selected channel and the transaction costs that have to be incurred in order to complete an exchange in the channel. The extension is necessary since the differences in transactions costs across different market channels, as well as the differential access to assets and services to assuage the transactions costs are likely factors underlying heterogeneous market participation decisions by smallholders (Alene *et al*., 2008).

Since the extension of the full income constraint creates discontinuities in the Lagrangian, the optimal solution of the utility problem cannot be found by simply solving the FONCs (Key *et. al.*, 2000). The solution is consequently decomposable into two steps, following Diane e*t. al.* (2003). In the first step, FONCs are used to derive the optimal standard input and output demands, conditional on the market channel-participation regime. In the second step, the market regime that leads to the highest level of utility is chosen.

The choice of the optimal market regime involves comparing utilities from alternative market channels and is therefore discrete and not amenable to FONCs (Rosenzweig, 1980). The conditions defining the optimal market regime can be modeled using the random utility model (RUM) developed by Marschak (1960) and McFadden (2000). Within RUM; a household faced with two alternatives π and μ will choose an alternative π if the utility $V\pi$ of choosing π is greater than the utility V_{μ} of choosing alternative μ (McFadden, 2000). The comparisons can be expressed as;

$$V_{i\pi}f(Z_{i\pi}) \geq V_{i\mu}f(Z_{i\mu}) \qquad for\ \pi \neq \mu \qquad (2.28)$$

Where $Z_{i\pi}$ and $Z_{i\mu}$ are column vectors of observable and non-observable household characteristics that determine the value that the household attaches to the attributes of alternative π and μ. Equation 2.28 can be used to estimate the probability with which each alternative is chosen. This study applied the RUM to model the smallholder household participation in ATMC in the study area.

2.2.4 Models for Estimation of Treatment Effects

The term 'treatment effect' refers to the causal effect of a binary variable on an outcome variable of policy interest (Cameron and Trivedi, 2005). In this study, this is equivalent to the effect of household participation in ATMC on TE. To analyze the effects (efficiency effects of household participation in ATMC), two potential outcomes; the outcome with treatment, TE_{1i} and the outcome without treatment TE_{0i} can be delineated. Using these outcomes, the average treatment effects (ATE) and the average treatment effects on the treated (ATET) can be derived as shown in equation 2.29 and 2.30.

$$ATE = E(TE_{1i} - TE_{0i} \mid Z_i) \quad (2.29)$$

$$ATET = E(TE_{1i} - TE_{0i} \mid Z_i;\ D_i = 1) \quad (2.30)$$

Where, Z_i is a vector of household characteristics and D_i denotes the household participation status in the treatment program. Unlike ATE which simply describes the expected effect of treatment for an arbitrary household with characteristics Z_i; ATET measures the mean effect of those who actually participated in the program. The measure is therefore more relevant for evaluating the treatment effects of a program and is achieved by comparing the performance of the participating households with the performance the same households would have achieved without participation (Veerbek, 2012).

A critical issue in the estimation of ATET is how households are assigned into the treatment program. Under random assignment, as is usually the case in

experimental or quasi experimental studies, ATET can be obtained as the difference in the average outcomes between the participants and non-participants as shown in equation 2.31;

$$\mathrm{ATET} = \frac{1}{\mathrm{N_P}} \sum_{i=1}^{N} \mathrm{TE}_{1i} - \frac{1}{\mathrm{N_{PN}}} \sum_{i=1}^{N} \mathrm{TE}_{0i} \qquad (2.31)$$

Where, $\mathrm{N_P}$ and $\mathrm{N_{PN}}$ is the sample size for the participants and non-participants respectively. This approach (equation 2.31) is similar to estimating the coefficient of discrete treatment dummy variable in a regression model. The approach is however inappropriate for the estimation of ATET using observational data (as opposed to experimental data) since the assumption of random assignment does not hold. The problem is that the estimates of ATET based on equation 2.31 would be subject to endogeneity biases (Cameron and Trivedi, 2005). These biases have the potential to render the coefficient estimates from standard regressions causally un-interpretable (Clougherty and Duso, 2015). The main sources of endogeneity as observed by Wooldridge (2002) are simultaneity and omitted variables. The former arises when the selection process of the agents to be analysed represents an excluded variable that manifests in the error term and correlates with the endogenous choice and outcome variable (Antonakis *et. al.,* 2010).

In the face of endogeneity, two approaches; the parametric regression and the semi-parametric, Propensity Score Matching (PSM) models are commonly applied in the estimation of treatment effects (ATET) when using cross sectional data. A review of the methods is presented in section 2.2.4.1 and 2.2.4.2.

2.2.4.1 Parametric Regression Models

The parametric approach to the estimation of ATET includes the Heckman model proposed by Heckman (1979) and the Endogenous Switching Regression (ESR) model proposed by Lee (1978). The choice between the Heckman and ESR models depends on the nature of the selection problem being addressed. This is because selection-based endogeneity manifests in two main forms; sample-selection and self-selection biases (Clougherty and Duso, 2015). Endogeneity due to sample selection is associated with the use of a non-randomly selected sample to estimate causal relationships and arises when observational units make decisions such that a subset of the target population is not observed, or when a subset of the population is excluded by the analyst through censoring (Heckman, 1979). Self-selection endogeneity on the other hand arises when the decision to participate in a treatment is voluntary and individuals self-select themselves into either the treated or non-treated states (Clougherty and Duso, 2015). Self-selection is thus a problem of the treatment not being randomly assigned to the agents being analyzed in cross sectional data (Antonakis *et al.*, 2010).

To solve for endogeneity due to sample-selection, the Heckman model proposed by Heckman (1974) can be applied. The model assumes that sample selection is a sort of missing variable problem and is implemented in two steps. In the first step, the model employs a selection equation to predict selection into the treated group then creates a variable that captures the differences between the sample for which outcome is observed and the sample for which the outcome is not observed (Wooldridge, 2002). In the second step, the model introduces a bias correction

term which is also known as the Inverse Mill's Ratio (IMR) into the outcome equation (the treatment effect model) to control for the bias that would otherwise manifest in the error term. The Heckman procedure yields unbiased parameters of the treatment effects when sample-selection is present. Following Heckman (1979), the estimation of treatment effects would be calculated using the following specification.

$$\mathrm{E}\left(\mathrm{Y_i} \middle| X_i, D_i = 1\right) = \beta' X_i + \rho\sigma_\varepsilon \left(\frac{\phi(\mathrm{Z_i}\alpha)}{1 - \Phi(Z_i\alpha)} \right) \qquad (2.32)$$

Where, $\mathrm{Y_i}$ is the outcome variable, $\mathrm{X_i}$ a vector that represents the treatment variable and other covariates and $\mathrm{D_i}$ denotes the treatment status of the agent being analysed. The term on the right hand side of the equation shows the product between the covariance of the error term of the selection and outcome equation $\rho\sigma_\varepsilon$ and the IMR. The IMR measures the likelihood that an ith observation is in the subsample $(D_i = 1)$. The parameter β provides a consistent and unbiased estimator of the treatment effects of X on Y; while $\mathrm{Z_i}$ is a vector of covariates determining the likelihood of $(D_i = 1)$ and $\emptyset$ (.) and Φ (.) are the standard and cumulative density functions respectively.

The ESR model whose foundation is the Heckman framework is suitable for addressing self-selection biases (Lokshin and Sajaia, 2004). Although the biases can also be solved using the Instrumental Variable (IV) framework; the approach has limited application given the difficulty of finding a suitable IV when the potentially endogenous treatment variable is binary (Clougherty and Duso, 2015).

The other key weakness of the IV approach is that it assumes that the treatment effect is represented by a parallel shift of the outcome equation and only accounts for intercept effects (Shiferaw *et. al.,* 2014).

ESR relaxes the assumption that treatment effect is only represented by a parallel shift of the outcome equation and accounts for both intercept and slope effects (Carrasco, 2001; Shiferaw *et. al.,* 2014). The method also accounts for both the observed and unobserved heterogeneities. In its general form, ESR consists of a selection equation and two (2) treatment outcome equations; for the treated and non-treated observations as shown in equation 2.33 to 2.35.

$$\mathrm{D_i^*} = \alpha' Z_i + \varepsilon_i \qquad (2.33)$$

$$\mathrm{Y_i} = \beta' X_i + \mu_i (D_i = 1) \qquad (2.34)$$

$$\mathrm{Y_i} = \beta' X_i + \mu_i (D_i = 0) \qquad (2.35)$$

Where $\mathrm{D_i^*}$ is a binary indicator variable that is equal to unity if the agent receives treatment and zero otherwise, Z_i is a vector of factors influencing the decision to receive treatment, Y_i is the outcome variable conditional on the treatment, X_i is a vector of observed characteristics that affect the outcome, β and α are vectors of parameters, and μ_i and ε_i are the error terms. Given that the choice to receive treatment is endogenous, the error terms in equations 2.34 and 2.35 conditional on treatment, have nonzero expected values due to self-selectivity as shown in equation 2.36 and 2.37.

$$\mathrm{E}\left(\mu_i \middle| \mathrm{D} = 1\right) = \sigma_{\varepsilon,\mu}\left(\frac{\phi(\mathrm{Z_i}\alpha)}{1 - \Phi(Z_i\alpha)}\right) \neq 0 \qquad (2.36)$$

$$\mathrm{E}\left(\mu_i|\mathrm{D}=0\right)=\sigma_{\varepsilon,\mu}\left(\frac{\phi\left(\mathrm{Z_i}\alpha\right)}{1-\Phi(Z_i\alpha)}\right)\neq 0 \qquad (2.37)$$

Where $\sigma_{\varepsilon,\mu}$ is the covariance between the error term of selection and outcome equation and $\emptyset$ (.) and Φ (.) are the standard normal probability density function and the standard cumulative density function respectively.

In light of the presence of self-selection endogeneity, estimation of equation 2.34 and 2.35 using standard least squares would lead to inconsistent estimates. The problem is solved by introducing an additional regressor into the outcome equations as shown in equation 2.38 and 2.39.

$$\mathrm{Y_i}=\beta' X_i+\frac{\phi\left(\mathrm{Z_i}\alpha\right)}{1-\Phi(Z_i\alpha)}+\mu_i(D_i=1) \qquad (2.38)$$

$$\mathrm{Y_i}=\beta' X_i+\frac{\phi\left(\mathrm{Z_i}\alpha\right)}{1-\Phi(Z_i\alpha)}+\mu_i(D_i=0) \qquad (2.39)$$

The additional regressor $\left(\frac{\phi\left(\mathrm{Z_i}\alpha\right)}{1-\Phi(Z_i\alpha)}\right)$ to correct for endogeneity is known as the IMR and is a monotone decreasing function of the probability that a particular observation for the variable of interest is observed in the outcome equation (Clougherty and Duso, 2015). Using equations 2.38 and 2.39, the treatment effects (ATET) of X_i on Y_i can be obtained by calculating the difference in the conditional expectations of Y_i conditional on X_i and the treatment status $\left(D_i=1\right)$. This study applied the ESR model in the estimation of the effects of ATMC participation since the household decision regarding ATMC participation is voluntary and therefore subject to self-selection endogeneity.

2.2.4.2 The Semi Parametric Propensity Score Matching Model

The PSM model proposed by Rosenbaum and Rubin (1983) is an alternative non parametric approach to dealing with self-selection endogeneity. The model works by comparing a group of treated and non-treated subjects who are similar in all relevant pre-treatment characteristics, and then attributes the differences in the outcomes between the groups to the treatment (Awotide *et al.,* 2015). PSM is based on the assumption that each subject has two potential outcomes; one if the subject receives treatment and the other if the subject is not treated. By comparing treated with untreated subjects who have the same potential outcome, PSM ensures that the difference between the treated and untreated subjects is due to the treatment, since the outcomes in both cases would have been the same had the treated not received treatment.

The matching of the treated and untreated subjects to ensure that they have the same potential outcomes before comparison is achieved using propensity scores. As shown by Rosenbaum and Rubin (1983) subjects with the same propensity score have on average the same potential outcomes and therefore comparing them can yield unbiased estimates of the effect of treatment. The procedure in PSM starts by obtaining propensity scores for a sample using a binary Probit or Logit model (Caliendo and Kopeinig, 2008). Propensity scores show the conditional probability of assignment of an observation into a treatment status given a vector of observed covariates. Following Rosebaum and Rubin (1983), the propensity score can be estimated as;

$$P(S)_i = \Pr_i(D_i = 1 | Z_i) \quad (2.40)$$

Where $P(S)_i$ represents the propensity score, and $\Pr_i(D_i = 1 | Z_i)$ is the conditional probability of an observation being assigned into the treatment given its observed characteristics (Z_i). The calculated propensity scores are then used to match the observations from the treated group with their counterparts in the non-treated set using a suitable matching algorithm. While there are many ways of matching, the nearest neighbour matching (NNM) method is the most applied. The method compares the propensity score of each treated observation with the propensity scores of the non-treated observations and selects a matching counterpart that is closest to it (Caliendo and Kopeinig, 2008: Awotide *et al.*, 2015). Successful matching in PSM relies on two (2) key assumptions; the conditional independence and common support or overlap. The first assumption implies that the distribution of potential outcomes for the units being analysed is the same regardless of their treatment status. The former assumes that all the treated observations have a counterpart in the non-treated group (Verbeek, 2012).

In the final PSM step, the treatment effect (ATET) is estimated by calculating the difference in outcome between the treated and the untreated and the difference is attributed to the treatment as shown in equation 2.41.

$$ATET = E(Y_T / D = 1) - E(Y_{NT} / D = 1) \quad (2.41)$$

Where $E(Y_T / D = 1)$ is conditional mean of the outcome for the treated contingent on participation in the treatment and $E(Y_{NT} / D = 1)$ is the conditional mean of the

outcome for the non-treated conditional on participation. For this study, PSM was employed to check the robustness of the estimates from the ESR model which was used to estimate the effect of household participation in ATMC on TE in the smallholder tea subsector.

2.3 Empirical Literature

Various studies have been conducted to investigate efficiency in agricultural production. The empirical literature discussed in this section begins with studies that analyzed efficiency in tea production. This is followed by a review of studies which investigated the relationship between farm size and efficiency in tea production and other related agricultural enterprises. The section concludes with a review of empirical studies that analyzed market participation decisions in tea production.

2.3.1 Efficiency in Tea production

Kavoi *et al.* (2001) investigated the relative economic efficiency of smallholder tea farmers in the East and the West of the Rift Valley. Efficiency was measured as the annual gross margin (GM) of the tea enterprise; where farms with higher GM were considered more efficient than those with lower GM. The calculated annual GM for each farm in the sample was then regressed against a dummy variable representing the two regions, the number of tea bushes and intensity in use of labour and fertilizer per hectare per year. The results revealed that farmers in the East of Rift Valley were more efficient than farmers in the West of the Rift Valley. In addition, the other explanatory variables in the model were found to have a positive and significant influence on efficiency among the sample farmers. The

study is important in elucidating some of the factors that would affect efficiency in tea production and in showing the contribution of regional differences on efficiency. The study is however faulted for using a measure of efficiency which does not correspond to the theoretical conceptualization of efficiency. Consequently, the study did not decompose economic efficiency into its theoretical components (TE and allocative efficiency).

Basnayake and Gunaratne (2002) estimated the TE in smallholder tea production in Sri Lanka using survey data of smallholder farmers. TE was estimated using SFM based on the Cobb-Douglas and translog specifications. The production inputs included in the Cobb-Douglas model as explanatory variables were land area under tea, family labour, hired labour, chemicals and dolomite. The translog model contained all the inputs in the Cobb-Douglas specification but also included their quadratic and cross- interactive terms. The study found that the mean TE was about 61 percent for the Cobb-Douglas model and 83 percent for the translog specification. The results of a model fitted to explore the sources of inefficiency showed that while the variety of tea, age, education, occupation and experience of the farmer had significant coefficients under the Cobb-Douglas specification; none of the parameters in the translog specification had significant effects. This study is important in providing insights on the determinants of TE in tea production. In addition, the differences in the results obtained from the two models appear to affirm the observation that SFM approaches are highly sensitive to specification. The present study adopted the non-parametric DEA model which is less prone to errors of misspecification in the estimation of TE.

Mwaura *et. al.* (2008) assessed the determinants of tea productivity using survey data of 208 smallholder farmers affiliated to Imenti and Kapkoros tea factories in Meru and Bomet counties of Kenya. In the study, productivity was measured as the output of tea per hectare and was regressed on various variables predicted to have effects on productivity. The results of the regression model revealed that the size of the household had a negative effect on productivity while the quantity of fertilizer applied, the proportion of owned land allocated to tea, the number of tea clones propagated per farm and access to extension services had a positive influence. Although this study provided some insights on the determinants of productivity in the tea sector, the measure of productivity which was used (output of tea per acre) is considered inappropriate since it gives too much importance to one input, land at the expense of the other inputs (Masterson, 2007). In addition, despite rising concerns about subdivision of tea farms, the study failed to explicitly consider the effects of farm size on tea productivity.

Kiprono (2013) used the SFM to compare the resource use efficiency of smallholder tea farmers in Kericho and Kiambu counties in the West and East of the Rift Valley respectively. Data was obtained from a sample of 384 tea farming households. The study found that smallholder tea farmers in both counties were inefficient in tea production and affirmed the finding by Kavoi *et. al.* (2001) that farmers in the East of Rift Valley were more efficient than farmers in West of the Rift Valley. The regional differences in efficiency were attributed to the differences in socio-cultural characteristics of farmers in the two counties. Although the study provides insights on the influence of locational heterogeneities

on efficiency, it did not account for the presence of geo-climatic differences that characterize the two study counties. The study is further faulted for its failure to incorporate the influence of market reforms in its models. The present study analyzed the effects of ATMC participation which is one of the key market reforms outcomes and controlled for geo-climatic variability by considering research sites with similarities in agro-ecological conditions.

Nguyen-Van and To-The (2014) used a translog SFM to investigate the effect of agricultural extension on TE of smallholder tea producers in northern Vietnam. Data was collected through a survey of a random sample of 244 smallholder tea farming households comprising of 130 green and 114 black tea producers. In the study, TE was estimated using a frontier model of the form;

$$\ln Y_i = \ln F(X_i, \beta) + v_i - u_i \qquad (2.41)$$

Where; Y_i was the quantity of harvested tea, X_i was a vector of inputs which included fertilizer, labour and land, while v_i and u_i were the error terms. To identify the sources of inefficiency, the TE scores from the SFM were regressed against the income and education level of the farmer, the variety of tea and various extension dummies representing different categories of extension information including cultivation techniques, fertilizer application, plucking, sales skills and market information. The study found that the mean farm size and annual output were 0.58 hectares and 4900 kgs, respectively. The mean TE was 0.323 and its distribution ranged between 0.005 and 0.915. With regard to extension effects, the model results showed that only the parameters of the dummies for sales skills and

market information were significant. The study concluded that adoption of appropriate extension models which put emphasis on marketing skills would enhance TE and demonstrated that institutional arrangements are important determinants of TE in tea production. The study is however faulted for not controlling for endogeneity in the estimation of extension effects since extension was potentially endogenous in the model which was estimated. The other variables that were found to have important influence on TE in the study were the variety and the type of tea, with black tea production being more efficient than green tea.

Hong and Yabe (2015) examined the resource use efficiency of tea production using a cross sectional data set of tea farming households in the Northern Mountainous Region of Vietnam. Efficiency was estimated using a translog model specified as;

$$\ln Y_i = \beta_0 + \ln \sum_j \beta_j \ln X_{ij} + \tfrac{1}{2} \sum_j \sum_k \beta_{ij} \ln X_{ij} X_{ik} + v_i - u_i \qquad (2.42)$$

$$j = 1...5, k = 1...5, \beta_{jk} = \beta_{jk}$$

Where; Y_i represents the output of the i^{th} tea farm, X_{ij} is vector of inputs that included fertilizer, pesticide, labour, irrigation water and capital, while v_i and u_i were the error terms. Results of the model showed that the mean level of TE was 0.82 which implied that farmers in the sample would reduce their observed levels of inputs by about 18 percent without causing a reduction in output. The mean elasticity of output with respect to the five (5) inputs was 0.323; which suggested presence of DRS in the production technology. A Tobit model estimated to

identify the determinants of TE revealed that enhancing the adoption of soil and water conservation technologies, extension services and household agricultural income would increase resource use efficiency in tea production. Results also showed that tea farms operated by women farmers were less efficient than those operated by males. The parameters of other variables in the model such as farm size, ethnicity, education, age of tea and experience were insignificant. Even though the study contributes to the existing literature on efficiency and its determinants; it is faulted for using the Tobit model which would have caused biases in the results.

2.3.2 Farm Size and Efficiency in Agricultural Production

Many studies that have investigated the effect of farm size on agricultural production have been based on partial measures of productivity such as yield per unit area (Balla and Roy, 1988, Gunther, 1995; Githinji *et. al*, 2011). The review in this section is based on efficiency measures which are broader and preferable since the partial productivity measures have been shown to be biased in favor of small producers.

Gilligan (1998) investigated the relationship between farm size and efficiency of coffee farms in in the Comayagua and Santa Barbara regions of Honduras. Data for the study was collected from a sample of 409 farms and efficiency was estimated using the non-parametric DEA model under assumptions of variable returns to scale. The effect of farm size on efficiency was estimated using the censored Tobit regression model which included farm size and land quality as explanatory variables. The study found an inverse relationship between farm size

and efficiency and made recommendations to support land reforms in favour of reductions in farm size. A decomposition of efficiency into technical and scale efficiency demonstrated that the production technology of the farms in the sample was dominated by scale diseconomies which therefore partly explained the observed inverse relationship.

Although the study provides some insights on what might explain the farm size-efficiency relationship, the use of the Tobit model is questionable since its domain is conceptually inconsistent with the nature of DEA scores. The argument is that even though the efficiency scores are bounded at the unit interval [0, 1], this is a consequence of the way DEA scores are defined rather than the result of censoring (Simar and Wilson, 2007).The present study employed the fractional regression (FR) model which accounts for the fractional nature of efficiency scores.

Helfand and Levine (2004) explored the determinants of TE and the relationship between farm size and efficiency using agricultural census data in Western Brazil. The study used DEA under CRS assumptions to estimate TE and regressed efficiency on a set of explanatory variables using generalized least squares (GLS) to take care of heteroskedasticity. The explanatory variables of the model were farm size, land tenure, composition of output based on the share of permanent, temporary and horticultural crops, access to institutions and public goods and the level of technology. The study incorporated a quadratic term for farm size to capture the influence of non-linearities and demonstrated a pioneering result that the effect of farm size on efficiency was quadratic with efficiency first dropping and then increasing with size. The study is important in showing the significance

of non-linear effects and therefore informed the inclusion of a quadratic term in the investigation of the effect of farm size on TE in the present study. The other variables that were found to be important in explaining efficiency differences across farms were the type of tenure, use of modern inputs and access to institutions and markets.

Rios and Shively (2005) used DEA to estimate the technical and cost efficiency of small coffee farms in Vietnam using survey data collected in two coffee producing districts of Vietnam. Using the DEA scores, Tobit regressions were carried out to show the effect of farm size on technical and cost efficiency (CE). Farm size in the study was captured by a dummy variable representing large and small farms, where farms with more than 1.5 hectares of coffee were classified as large farms while those with 1.5 hectares or less were categorized as being small. The other explanatory variables were the education level of the household head, the type of land tenure arrangement, irrigation infrastructure and access to credit. To account for possible endogeneity in the credit variable, the model was instrumented using the inverse mills ratio (IMR) derived from a probit model as an additional regressor.

The results of the instrumented Tobit showed that contrary to the frequently acclaimed inverse-efficiency hypothesis, the small farms were less efficient than the big farms. The finding was attributed in part to the scale of investment in irrigation infrastructure, indivisibilities in certain technologies which favoured the large farms and the fact that the large farms had preferential access to inputs and institutions. The study contributes to the literature on farm size and efficiency and

provides insights on the possible determinants of the farm size-efficiency relationship. Apart from the use of the Tobit, the study is also faulted for adopting a very subjective classification of farm size which would not allow the pinning down of a specific threshold of farm size that would be targeted for farm size reforms.

Masterson (2007) used a data set of 8131 agricultural households in Paraguay to assess the relationship between farm size and productivity and TE in Paraguay. Productivity was captured as the aggregate value of agricultural output per unit area while TE was estimated using both the SFM and the LP DEA model. After the estimation of TE, Tobit and OLS regressions were estimated to respectively identify the determinants of efficiency and productivity in the sample. The study found that all the three (3) coefficients of farm size in the OLS, SFM Tobit and DEA Tobit estimations were negative. The parameters were again negative in the case of the square of farm size for the OLS and SFM Tobit specifications, but positive under the DEA Tobit. The study concluded that relationship between farm size and efficiency was inverse and recommended measures to bolster redistribution of land in Paraguay.

Contrary to the labor market-imperfections hypothesis which appears to dominate the farm size-productivity literature, the study found that rising shares of family labor employed in farming resulted in lower levels of efficiency. Apart from contributing to the farm size- productivity debate, the study by Masterson is important in pointing out the possible biases that would arise when partial measurers of productivity are applied in the investigation of the relationship

between farm size and productivity. Additionally, the two step approach applied in Rios and Shively (2005) and Masterson (2007) was adopted in the present study in the analysis of efficiency effects of farm size in the smallholder tea subsector in Kenya.

Using survey data of 309 smallholder tea farmers in Kenya, Ogise *et. al.* (2008) conducted a study to determine the minimum threshold of farm size required for economic tea farming. The study used a modified cost-return model in which the net returns from tea farming were obtained by deducting the costs of farm operations and household necessities from the gross tea earnings. The minimum land requirement was then calculated as the tea acreage that would meet the cost of farm activities and annual household annual expenditure. The study established that tea farming was a key source of income in the tea growing areas and showed that most of the households (77 percent) in the sample were operating below the minimum threshold. The study also found that the minimum threshold varied across households due to differences in production costs and productivity. The study recommended that apart from increasing the area under tea, it was necessary to strengthen measures to improve tea productivity. While the study is important in highlighting the significance of farm size issues in Kenya's tea sector; there is need to revisit the subject in order to determine an efficiency-based threshold of farm size that would guide farm size reforms in the sector.

Padilla-Fernandez and Nuthall (2012) used survey data and the DEA model to examine the effect of farm size on the productive efficiency of sugar cane production in Philippines. The study was conducted against the backdrop of

controversies on whether the agrarian policy reforms in the country that restricted ownership of land to 5 hectares were enhancing or constraining efficiency in agriculture. The study found that small farms were not as efficient as the larger farms and that the medium and large farms were equally efficient. The study concluded that policy reforms promoting larger farm sizes would have a beneficial impact on efficiency. The study attributed its result to the fact that smaller farms faced higher input prices which reduced their usage.

2.3.3 Household Participation in Agricultural Markets

The participation in agricultural markets is arguably an essential precondition for agricultural growth (Fafchamps, 1992). Consequently, empirical analysis on the household participation in agricultural markets is highly acclaimed in literature (Goetz, 1992; Govereh *et. al.,* 1999). A common theme in the literature is that transactions costs can exert strong influence on farmers' decisions to participate in markets (Fafchamps and Hill, 2005; Omiti *et. al.*, 2009). While this conclusion is relevant for crops that can either be sold or consumed at home, the aspects of market participation that relate to tea farming are relatively different. Since all the tea produced by farmers is intended for sale, the relevant market participation issue is not whether households participate in markets or not but rather the choice of market channels through which they participate. The review in this section is focused on the choice of tea market channels for the tea farming households.

Chirwa and Kydd (2006) carried out a study to investigate the effect of contractual and marketing arrangements on smallholder tea productivity in Malawi. The study used a multivariate linear regression model, where participation in green leaf

market channels, the intensity of input use, a vector of farmer characteristics which included the age, gender and education of the household head and farm size were included as explanatory variables. Findings from the study revealed that farmers who had contracts with commercial estates produced more yield than those who had contractual arrangements with the reformed state-enterprise factories. This was attributed to the differential services offered by the factories, with the commercial estates offering superior extension services in addition to input credit and timely supply of inputs. Although the study's model did not control for self-selection endogeneity, it is useful to the extent that it demonstrates that the type of market channel through which farmers sell their tea output is an important determinant of productivity in tea production.

Using survey data from 245 smallholder tea farmers, Yoshiko (2011) analysed the effect of institutional and marketing arrangements on TE and income in the Moc Chau District of Vietnam. The institutional arrangements that the study considered were; farmers contracted with a state-owned enterprise (SOE), contracting with a private organization and the non-contracted farmers. In order to account for the sample selection biases, the study employed the non-parametric propensity score matching (PSM) model. The study found that smallholder farmers contracted with the SOE had significantly higher levels of TE than both the non-contracted farms and those contracted with a private enterprise. The study concluded that better performance of the SOE was due to its precise and reliable extension services. The present study adopted both the parametric endogenous switching regression (ESR)

and the non-parametric PSM model to assess the effect of participation in alternative market systems on TE in the tea sector in Kenya.

Chepkulei (2013) investigated the factors that influence choice of a tea marketing channel among smallholder tea households in Kenya. Data for the study was collected from a survey of 182 households and was analysed using the Logit regression model. The model, considered the binary choice whether a household was selling its produce through KTDA or otherwise (private company or middleman). The regression results revealed that institutional factors mainly promptness in payment, green leaf quality requirements and flexibility in collection of green tea favorably predicted a farmer's decision to sell to the alternative channels. Additionally the study found that farmers who were participating in the alternative markets were generally younger; more educated and had smaller household sizes. While Chepkulei's study provides useful insights on the determinants of ATMC participation, it did not consider the effect of ATMC participation on efficiency and therefore whether participation in these alternative channels impacts TE in the Kenyan tea sector remains unknown.

2.4 Overview of Literature

The theoretical literature reviewed in this study is grounded on the agricultural production theory and the theory of the utility maximizing peasant. The review of the agricultural production theory provided information on the characterization of agricultural production technology, the conceptualization of TE and the theoretical basis for modeling smallholder production and estimation of TE. The review of the theory of the utility maximizing peasant provided the theoretical underpinnings for

modeling ATMC participation by households. A part from the review of the theories, the theoretical literature showed that the SFM and DEA are the most common approaches for estimation of TE and that while the SFM has capability to separate inefficiency from statistical noise; DEA is less prone to misspecification and provides a fuller characterization of technology.

The review of empirical literature shows that studies that analysed efficiency in agricultural production have used both DEA and SFM models. The reviewed studies provide useful insights on the characterization of agricultural production technology and the determinants of efficiency in tea production. In addition, the review shows that previous studies on efficiency on tea production in Kenyan have been undertaken in the context of regional comparisons of efficiency between the East and West of Rift Valley. Although the studies highlight the contribution of locational heterogeneities in explaining efficiency, they are faulted for their failure to control for the geo-climatic variability that characterizes the two regions.

With regard to the relationship between farm size and efficiency, the review shows that while the subject has attracted a lot of interest in other countries; the issue has received limited research attention in Kenya. Most of the studies that have analysed the effect of farm size on efficiency have focused on other industrial crops such as coffee and were also based in other countries. Consequently, despite the fact that farm size and sub-division of tea farms has elicited a lot of concerns from stakeholders, there is lack of empirical evidence to guide and support farm size reforms in Kenya's tea sector.

Finally, the review has demonstrated the lack of empirical evidence linking efficiency with the liberalization reforms which were introduced to spur efficiency in the Kenyan tea sector. In light of the foregoing and the need to revisit the analysis of efficiency due to the changes that have taken place; this study undertook a more comprehensive analysis of TE and incorporated the effects of farm size and marketing arrangements in its models. The study also ascertained the influence of other farm and household as well as policy and institutional variables on TE. The research sites which were selected for the study have similar agro-ecological conditions which minimized the effects of geo-climatic variability in the analysis of efficiency.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter discusses the research design and the theoretical and empirical models applied in the analysis of efficiency and the effects of farm size and green leaf marketing on technical efficiency (TE). A description of the variables and their measurement, data collection approach, target population, sampling procedure and sample size and data analysis is also provided.

3.2 Research Design

The study employed a non-experimental cross-sectional research design to achieve its objectives. The design was considered suitable since the random assignment of observations which is necessary in experimental research designs was not feasible in the case of farm households. The design allowed for the collection of information on household farm, demographic, socioeconomic and institutional characteristics that was necessary for the analysis of efficiency. Data for the study was collected from a cross sectional survey of 525 smallholder farmers from Bomet and Nyamira Counties of Kenya.

3.3 Theoretical Framework

3.3.1 Theoretical Framework on Measurement of Technical Efficiency

The estimation of TE in this study follows a framework based on agricultural production theory in which a typical tea farming household is assumed to use owned and purchased inputs to produce tea and other farm outputs. The

household's production technology therefore utilises a vector of inputs denoted $X = (x_1, \dots, x_n) \in \Re_+^n$ (Euclidean space) to produce a non-negative vector of output denoted $Y = (y_1, \dots, y_n) \in \Re_+^m$. The household's production possibility set (PPS) which is the collection of all the feasible input-output vectors is the sub set T of the space $\Re_+^{m+n}$ and is represented as;

$$T = \{(YX): X \text{ can produce } Y\} \in \Re_+^{m+n} \quad (3.1)$$

The inputs and outputs are assumed to be freely disposable, and T is assumed to be non-empty, closed, and convex. Following Fare, Grosskopf and Lovell (1994), the PPS can be represented by an input requirement set L(y) or an output producible set P(x) as shown in equation 3.2 and 3.3 respectively.

$$L(y) = \{X: (XY) \in T\} \quad (3.2)$$

$$P(X) = \{Y: (XY) \in T\} \quad (3.3)$$

The input requirement set represents the collection of all input vectors $X = (x_1, \dots, x_n) \in \Re_+^n$ that yield at least the output vector $Y = (y_1, \dots, y_n) \in \Re_+^m$. The output producible set is the collection of all the output vectors $Y = (y_1, \dots, y_n) \in \Re_+^m$ that can be produced from the given input vector $X = (x_1, \dots, x_n) \in \Re_+^n$.

The tea farming household may therefore select any input-output configuration $(XY) \in T$ as it's production plan and faces an optimization problem (since it is assumed to be profit maximizing) based on inputs and outputs. Analysis of the performance of a tea farm therefore requires specification of a technical relationship that reflects the choice of a production technology and combinations

of inputs that leave the farm with the most output given its feasible output set (Varian, 1992). According to Farell (1957), TE may be examined from either an input-orientation or output-orientation. In the input-orientation, TE means that a farm minimizes the quantity of inputs used in the production of a given level of output. The output-orientation on the other hand focuses on maximizing output from a fixed set of inputs.

Assuming conditions of VRS, the production technology of a firm can be represented on an input- output space as shown in figure 3.1.

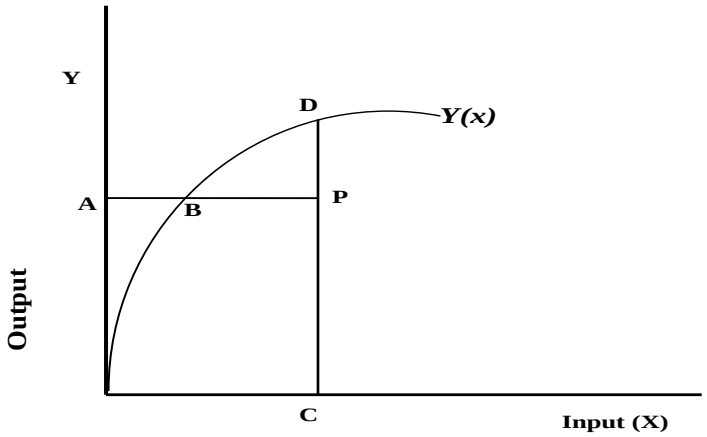


Figure 3.1: Input and Output Oriented Measures of TE

In figure 3.1, the frontier Y(x) is assumed to be an efficient production function. All points below the frontier such as point P; represents inefficient production because it is possible to increase production from P to D with the same level of input.

From an output orientation, the TE score for a farm producing output level P with input level C in the figure 3.1, can be calculated as;

$$TE = {}^{CP}/_{CD} \quad (3.4)$$

where CP is the production level of a given farm and CD represents the production level of a technically efficient or a frontier farm given input level C. In this case, the TE score tells us the level of success of the farm in producing maximum output from the fixed set of inputs.

From an input orientation, the TE score of a farm would be calculated as;

$$TE = AB/AP \quad (3.5)$$

where AB represents the input level a technically efficient (frontier) farm uses in producing the quantity of output equivalent to that which the given farm produces using AP units of inputs. In this case, the TE score enables the determination of the amount by which inputs could be proportionally reduced without a reduction in the farm's output. A tea farm is therefore considered efficient if its output is the maximum that can be produced using all of its inputs, but is inefficient if its output can be produced by a best practice farm using only a fraction of its inputs.

Since tea is a perennial crop and the output quantities and prices are rarely known with certainty when farmers make production decisions, it is more reasonable to assume that the tea production process starts with a fixed set of resources, rather than with a predetermined level of output (Coelli, 1995). The problem of the household is therefore to choose a production plan that will produce the highest level of output from a given set of resources devoted to tea production. The farm will be considered efficient if there exists no other production method that can use

less inputs than its current level to realize its level of output. The level of the TE score in the input orientation approach can be estimated by use of DEA.

3.3.2 Sources of Inter Household Variation in TE

The factors that influence efficiency variation across farms may be summarized into agent and structural factors (Ogada *et. al.,* 2014). These may consist of policy and institutional variables, the internal structure of the farm and agency factors such as the levels of human capital and experience of the farmer (Yoshiko, 2011; Kiprono, 2013). Therefore the TE of a farm can be conceptualized to be dependent on various factors as shown in figure 3.2.

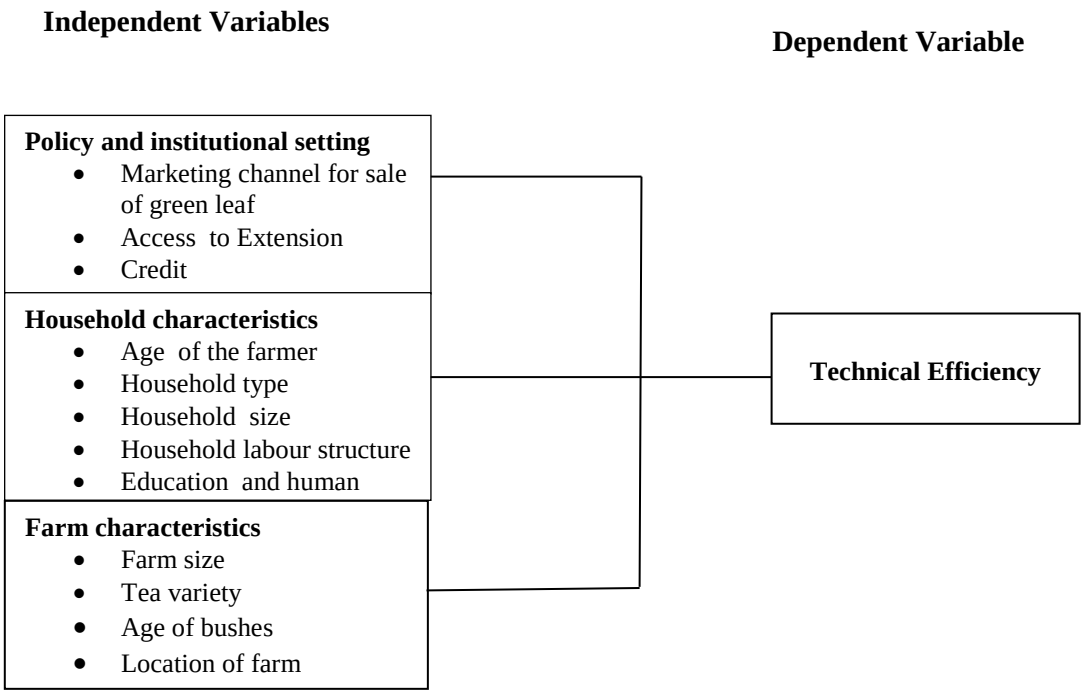


Figure 3.2: A Conceptualization of the Determinants of TE

Figure 3.2 shows the various marketing and institutional as well as farm level and household variables that are hypothesized to influence the level of TE among tea farmers. Using the efficiency scores obtained from DEA, it is possible to

determine how the variables influence TE across farms by applying the Fractional Regression (FR) Model proposed by Papke and Wooldridge (1996) to model dependent variables which are defined within the unit interval (Wooldridge 2010; McDonald 2009; Ramalho *et. al.,* 2010). The model is appropriate for analysing DEA scores since it is based on a functional form that imposes the restriction that the conditional mean of the dependent variable is bounded within the interval [0, 1] as shown in equation 3.6;

$$\mathrm{E}(\mathrm{TE}\,|\mathrm{X}) = g(\mathrm{X}\beta) \qquad 3.6$$

Where, the vector X contains farm size as the variable of interest and other control variables hypothesized to be correlated with TE and $g(.)$ is some nonlinear function satisfying the condition that $0 \le g(z) \le 1$ for all $z \in R$ (Papke and Wooldridge, 1996). The function $g(.)$ is typically modelled as a cumulative distribution function and is commonly estimated as a logistic or standard normal distribution functions (Wooldridge 2002; Ramalho *et. al.,* 2011). The model keeps the predicted values in the unit interval through a more refined and flexible analysis and therefore represents a more natural approach to modelling bounded and fractional DEA scores (Ramalho *et. al.,* 2010; Gelan and Muriithi, 2015).

Although it is possible to keep the predicted values of TE within the unit interval using log–odds transformation $\left(\log(\mathrm{TE}/1-\mathrm{TE})\right)$ which allows use of OLS, the approach cannot be applied directly when the scores take on boundary values. While adjustments can be done on the boundary values, such adjustments would

necessarily be arbitrary (Wooldridge, 2002). The standard linear model (SLM) which has been applied in some studies was therefore not used in this study.

The other primary reason why SLM would not be used is that since the DEA scores are bounded, the effect of any particular covariate cannot be expected to be constant throughout the range of the covariates X (Papke and Wooldridge, 1996). Consequently the constant marginal effect of the covariates on TE implied by the linear conditional mean model is not consistent with both the bounded nature of the TE scores and the existence of a mass point at unity in their distribution (Ramalho *et. al.*, 2010).

The use of the two-limit Tobit model with limits at zero and unity was also not applied since the bounded nature of the DEA scores is a consequence of the way DEA scores are defined rather than the result of censoring (Simar and Wilson, 2007). Additionally, the domain of the Tobit model differs from that of DEA scores because typically efficiency scores of zero are not observed implying therefore that the Tobit model is an inconsistent estimator (McDonald 2009; Ramalho *et. al.,* 2010). Additionally, the model is highly sensitive to its assumptions on normality and homoscedasticity.

3.3.3 Theoretical Framework on the Effect of Marketing on TE

In order to analyse the effect of green leaf tea marketing on the TE, the study considers that a household can either sell its green leaf exclusively through the KTDA system or to an alternative market channel (ATMC). In the latter case, the

tea farming household may sell its green leaf through a middleman at the farm gate, or roadside or directly to a private factory.

Following equation 2.28 in chapter two, the household's choice of a marketing channel can be modelled using a random utility formulation (McFadden, 1973). This is because the decision to participate in a particular channel is not random but dependent on expectations of the household based on the costs and benefits of participation (McFadden, 2000).The farm household is therefore assumed to maximize its payoff when choosing a tea marketing channel. If U_p is the net benefit stream from ATMC participation and U_{pn} is the net payoff due to non-participation, a tea farming household would choose ATMC participation if doing so would leave it better than not participating as shown in equation 3.7.

$$D_i^* = U_p - U_{pn} > 0 \qquad (3.7)$$

Where, D_i^* is a latent variable that reflects the differences in the net benefits (payoffs) due to the household participation in ATMC and net payoffs due to non-participation in ATMC. The binary variable (D_i^*) is determined by observed characteristics(Z_i) and is equal to 1 if the household participated in an ATMC and zero otherwise as shown in equation 3.8

$$D_i^* = \alpha' Z_i + \varepsilon_i \qquad with\ D_i = \begin{Bmatrix} 1\ if\ D_i^* > 0 \\ 0\ othersise \end{Bmatrix} \qquad (3.8)$$

Where Z_i is vector of characteristics that affect the choice of a marketing channel, α is a vector of parameters and ε_i is the error term. The binary model underlying

the specification in 3.8 depends on the distribution of ε_i. The study assumed that ε_i followed the standard normal distribution which yielded the probit model.

In order to derive the efficiency effects due to ATMC participation, two separate TE equations conditional on ATMC participation were specified. The equations are written as an endogenous switching regime model as shown in equations 3.9 and 3.10.

$$\mathrm{TE_P} = X\beta + \mu_1 \qquad \text{if} \quad D_i = 1 \qquad (3.9)$$

$$\mathrm{TE_{NP}} = X\beta + \mu_0 \qquad \text{if} \quad D_i = 0 \qquad (3.10)$$

Where X is a vector of explanatory variables hypothesised to influence efficiency, β and α are the unknown parameters and μ_1 and μ_0 are the error terms. The error terms (in equation 3.8 to 3.10) are assumed to have a tri-variate normal distribution with mean vector zero.

Since the decision to participate in ATMC is voluntary and non-random, the switching model in equation 3.9 and 3.10 is prone to endogeneity and self-selection biases due to unobserved farm and household characteristics or heterogeneities (Lokshin and Sajaia, 2004). Following Maddala, (1986), correction of the biases can be achieved using the inverse mills ratio (IMR) which is computed from the binary model as shown in equation 3.11 and 3.12.

$$\mathrm{E}(\mu_1|\mathrm{D_i} = 1) = \sigma_{\varepsilon 1}\frac{\emptyset(Z_i\alpha)}{1-\Phi(Z_i\alpha)} = \qquad \sigma_{\varepsilon 1}\gamma_{i1} \qquad (3.11)$$

$$\mathrm{E}(\mu_0|\mathrm{D_i} = 0) = -\sigma_{\varepsilon 2}\frac{\emptyset(Z_i\alpha)}{1-\Phi(Z_i\alpha)} = \qquad \sigma_{\varepsilon 2}\gamma_{i2} \qquad (3.12)$$

Where $\sigma_{\varepsilon 1}$ is the covariance of ε_i and μ_1 and $\sigma_{\varepsilon 2}$ is a covariance of ε_i and μ_0 while $\emptyset$ (.) and Φ (.) are the standard normal probability density function and the standard cumulative density function respectively. The terms γ_{i1} and γ_{i2} are the IMRs that are used to correct the selectivity bias in switching regime model as shown in equation 3.13 and 3.14. Inclusion of the IMR as an additional regressor takes care of the unobserved variables and therefore corrects for the endogeneity or selectivity bias (Guo and Fraser 2010). The IMR reflects the probability that an observation belongs to the selected sample (Heckman, 1979).

$$TE_P = X\beta + \sigma_{\varepsilon 1}\gamma_{i1} + \mu_1 \quad if\ D_i = 1 \qquad (3.13)$$

$$TE_{np} = X\alpha + \sigma_{\varepsilon 2}\gamma_{i2} + \mu_0 \quad if\ D_i = 0 \qquad (3.14)$$

Using the estimations from equations 3.13 and 3.14, the efficiency effects of ATMC participation (E_{ATMC}) were derived by obtaining the difference in the expectations of TE conditional on ATMC participation (equation 3.15). The difference shows the change in TE due to ATMC participation and is known as the average treatment effects of the treated (ATET) in treatment effects literature.

$$E_{ATMC} = E(TE_p | D = 1; X) - E(TE_{pn} | D = 1; X) \qquad 3.15$$

Equation 3.15 calculates the difference between the participation (treatment) effects for the participants given the covariates X and the non-treatment (non-participation) effects for the participants conditional on the covariates X. The equation is based on the conditional independence assumption which states that given a set of covariates X; the potential non treatment effects are independent of the participation status as shown in equation 3.16 (Verbeek, 2012).

$$E(TE_{pn}|D=1;X) = E(TE_{pn}|D=0;X) \quad 3.16$$

The conditional independence assumption therefore allows for the determination of the non-participation effects (how the participant farms would have performed had they not participated) in equation 3.15.

3.4 Specification of Empirical Models

Three empirical models were specified to address the three specific objectives of the study. The models are presented in the following sections.

3.4.1 Estimation of Technical Efficiency

Following the theoretical framework described in section 3.3.1, The TE of smallholder farms was estimated using the DEA model under VRS (equation 3.5). This was necessary to allow the determination of scale inefficiencies. The linear programming VRS DEA model was specified as follows;

$$minimize_{\theta\,\lambda}\ \theta \quad (3.17)$$

Subject to;

$$-y_q + (y_1\lambda_1 + y_2\lambda_2 + \cdots + y_n\lambda_n) \geq 0 \quad (3.18)$$

$$\theta x_1^q - (x_{11}\lambda_1 + x_{12}\lambda_2 + \cdots + x_{1n}\lambda_n) \geq 0 \quad (3.19)$$

$$\theta x_2^q - (x_{21}\lambda_1 + x_{22}\lambda_2 + \cdots + x_{2n}\lambda_n) \geq 0 \quad (3.20)$$

$$\theta x_3^q - (x_{31}\lambda_1 + x_{32}\lambda_2 + \cdots + x_{3n}\lambda_n) \geq 0 \quad (3.21)$$

$$\lambda \geq 0\ (\lambda = \lambda_1, \lambda_2, \ldots, \lambda_n) \quad (3.22)$$

$$\sum_{n=1}^{n} \lambda = 1 \quad (3.23)$$

Where θ is the value of the efficiency score for the q^{th} farm and represents the proportion of the farm's input bundle needed to produce its own output, y_q denotes

the tea output of the q^{th} farm, y_i *(i = 1, 2, …, N)* refers to tea output of the farms in the sample, $\boldsymbol{x_i^q}$ **(** *i = 1, 2, 3)* denotes the level of the i^{th} input used on q^{th} farm (x_1 is the fertilizer, x_2 is the labour and x_3 is the land used in tea production) and λ is the weight given to each farm in the construction of the frontier. The model is interpreted as seeking a frontier farm that can produce at least the output of the q^{th} farm, using the smallest possible multiple of its inputs. The model was solved N (sample size) times to obtain efficiency scores for all the farms in the sample.

The characterization of the nature of the returns to scale (RTS) was achieved by estimating an additional non-increasing returns to scale (NIRS) DEA model in which the convexity constraint in equation 3.23 was replaced with $\left(\sum_{n=1}^{n} \lambda \leq 1\right)$. The RTS of the individual tea farms in the sample was then determined by checking whether the TE scores obtained from the VRS and NIRS models were equal. As argued in Banker *et. al.* (2004), a farm is considered to experience decreasing or increasing returns to scale if the TE score obtained from NIRS and VRS DEA models are equal or not equal respectively.

The final step in the analysis of TE involved an extension of the VRS DEA model to account for input slacks as shown in equation 3.24 to 3.31. The slack values indicate the amount by which the constraints in the model are not fully satisfied and therefore represent the amount by which inputs are overused relative to the efficient farms (Padilla-Fernandez and Nuthall, 2012). The analysis of slacks was necessary since non-zero slacks can represent a substantial amount of inefficiency (Seiford and Zhu 1999) such that a firm is considered fully efficient when its DEA

score is equal to one and all its slacks are zero (Seiford and Zhu 1999; Morita *et. al.,* 2005). The non-zero slacks arise when sections of the piece wise linear frontier estimated by DEA run parallel to the axes (Coelli, 1996)

$$minimize_{\theta\,\lambda S^-}\ \theta \tag{3.24}$$

Subject to;

$$-y_q + (y_1\lambda_1 + y_2\lambda_2 + \cdots + y_n\lambda_n) \geq 0 \tag{3.25}$$

$$\theta x_1^q - (x_{11}\lambda_1 + x_{12}\lambda_2 + \cdots + x_{1n}\lambda_n) - s^- = 0 \tag{3.26}$$

$$\theta x_2^q - (x_{21}\lambda_1 + x_{22}\lambda_2 + \cdots + x_{2n}\lambda_n) - s^- = 0 \tag{3.27}$$

$$\theta x_3^q - (x_{31}\lambda_1 + x_{32}\lambda_2 + \cdots + x_{3n}\lambda_n) - s^- = 0 \tag{3.28}$$

$$\lambda \geq 0\ (\lambda = \lambda_1, \lambda_2, \ldots, \lambda_n) \tag{3.29}$$

$$S^- \geq 0 \tag{3.30}$$

$$\sum_{n=1}^{n} \lambda = 1 \tag{3.31}$$

Where S^- are the input slacks. The measure of efficiency based on this model is considered to be more comprehensive since it corresponds to the Pareto-Koopmans concept of efficiency (Pareto (1909), Koopmans (1951) which states that a firm is fully efficient if and only if it is not possible to improve any of its inputs or outputs without worsening some other of its input or output (Cooper *et. al.,* 2007).

3.4.2 Effects of Farm Size on TE

From equation 3.6, the following FR model (equation 3.32) was estimated to determine the effect of farm size on TE. The model was also used to ascertain the influence of other variables such as the household and farm characteristics on TE.

$$\mathrm{TE_i} = g(z)[\mathrm{X_j}\beta] + \varepsilon_i \quad \text{for i} = 1,2,\ldots,\mathrm{N} \text{ and } \mathrm{j} = 1,2,\ldots,11 \tag{3.32}$$

Where, TE_i represents the efficiency score obtained from DEA; $g(z) \equiv \phi(z)$ is the standard normal cumulative distribution function; X_j is vector of covariates that include farm size, square of farm size, county dummy, variety of tea, age of the tea farm, square age of the farm, gender of the household head, education, FFS extension (FFS), distance to markets, labour structure and market channel dummy; β is a vector of the unknown parameters to be estimated and ε is the error term which is assumed to be $NID(0, \sigma^2)$.

3.4.3 Effect of Marketing System on TE

Following equations 3.8 and 3.13 and 3.14 the endogenous switching model for evaluating the effects of marketing arrangements on TE is specified in equations 3.33 to 3.35. The model consists of a binary equation that links the ATMC participation decision with the possible covariates of participation and separate outcome equations for the participants and non-participants conditional on ATMC participation.

$$D_i^* = \alpha' Z_i + \varepsilon_i \qquad \text{with } D_i = \begin{Bmatrix} 1 \text{ if } D^* > 0 \\ 0 \text{ othersise} \end{Bmatrix} \tag{3.33}$$

$$TE_{1i} = \beta_{1i} X_{1i} + \sigma_{1i}\gamma_{i1} + \varepsilon_{1i} \qquad \text{if} \qquad D_i = 1 \tag{3.34}$$

$$TE_{2i} = \beta_{2i} X_{2i} + \sigma_{2i}\gamma_{i2} + \varepsilon_{2i} \qquad \text{if} \qquad D_i = 0 \tag{3.35}$$

In the model, D_i^* is a latent variable observed through the decision to participate in a market channel and α is a vector of unknown parameters to be estimated. Z_i is a vector of exogenous variables hypothesized to influence the participation decision and included various farm level characteristics (farm size, county dummy, variety of tea planted, age of farm, volume of tea harvested and distance to the nearest

market), the household demographic characteristics (age of farmer, education, household labour structure and gender and residence of the household head), the household socioeconomic characteristic (household assets and per capita expenditure) and institutional variables (access to FFS extension and credit). TE_{1i} and TE_{2i} are the separate efficiency functions for the participant and non-participant households respectively, β_{1i}, β_{2i}, are vectors of parameters to be estimated and X_{1i} and X_{2i} are vectors of technical efficiency determinants and included the explanatory variables specified in equation 3.32. Finally the terms γ_{i1} and γ_{i2} are the inverse mills ratios and ε_i, ε_{1i} and ε_{2i} are the error terms which are assumed to have a joint trivariate normal distribution, with zero mean.

3.5 Definition and Measurement of Variables

The definition and measurement of variables in the models that were estimated are outlined as follows;

a) **Output:** The quantity of harvested tea (Green leaf) in kilograms per year by a farm.

b) **Fertilizer:** The quantity of fertilizer applied on the tea farm per year measured in 50 kgs bags.

c) **Education:** The education status of the household head measured in terms of the highest level of education attained (1= primary, 2 = secondary, 3 = tertiary 4= university).

d) **Farm size:** This is the total area under tea in acres.

e) **Labour:** The quantity of labour applied in the performance of tea production activities measured in man-days per year, where one man-day is equivalent to 6 hours.

f) **Labour structure:** The proportion or percentage of family labour applied in tea farming.

g) **Age of farmer:** The age of the household head measured in years.

h) **Gender:** This is defined in terms of the gender of the household head and was measured using dummies; where 1 represents a male headed household and 0 otherwise.

i) **Extension services (FFS):** This is participation of the household in the farmer field school (FFS) program. The variable was measured using dummies; where 1 represents participation in FFS and 0 otherwise.

j) **Tea variety:** This is the type of tea cultivar planted on the farm and was measured in terms of 1 if the tea farm had clonal (improved tea) and 0 otherwise.

k) **Household assets:** The value of selected household assets (Appendix 3) owned by the household, measured in Kenya shillings.

l) **County dummy:** The variable was used to distinguish the two counties from which the tea households were sampled, where Bomet was equal to 1 and Nyamira County was equal 0.

m) **Age of the tea farm:** The age of the tea farm measured in number of years since current bushes were establishment or planted.

n) **Credit:** Describes receipt of borrowed money for tea farming activities measured using dummies; where the variable is equal to 1 if the household had borrowed for agricultural activities and 0 otherwise.

o) **Per capita expenditure:** The per capita amount of money spent on the purchase of a basket of selected household necessities (Appendix 3) per year measured in Kenya shillings.

p) **Food sufficiency:** The number of months the household relied on the market to access food staples and was measured in number of months and was used as proxy for household liquidity demand.

q) **Technical efficiency (TE):** The TE scores obtained from the DEA model

r) **Market Channel (ATMC participation):** This is participation of the household in an ATMC. The variable was measured using dummies; where 1 represents participation in ATMC and 0 otherwise.

s) **Residence:** Availability of the household head on full time basis to make tea farming decisions. The variable was measured using dummies; where 1 represents availability and 0 otherwise

3.6 Target Population and Sampling

The survey targeted a population of 528, 817 smallholder tea farming households spread across the tea growing counties in Kenya as shown in Table A.1 in the Appendix. Based on the following statistical formula for determination of sample size proposed by Kothari (2004), the study interviewed a total of 525 tea farming households in Kenya.

$$N = \frac{Z^2 PQ}{d^2} \qquad 3.36$$

Where N is the desired sample size, Z is the standard normal distribution value which corresponds to 1.96 at 95% confidence level, *P (P=0.3)* is the proportion of

population with characteristics of interest, *Q* is *1-P* and *d* is the allowable margin of error which is assumed not to exceed 5% as done in many studies. The P value of 0.3 represents the proportion of tea growing counties where the alternative green leaf marketing system operates and is also a good value to use in a normal rural livelihood settlement (Kiprono, 2013; Republic of Kenya, 2014b).

The study adopted the multistage sampling procedure. The first stage involved a purposive selection of two tea growing counties; Bomet and Nyamira County. The counties were selected because the ATMC system which was under investigation in this study greatly thrives in them and they are among the leading tea producing counties in Kenya where there exists variability in tea farm sizes (KTDA, 2013). Further, the scale of farm holding in the two counties has representation of both the small and relatively large scale tea farms. A profile of the two study counties is described in Appendix 2.

The second stage involved the stratification of the counties into administrative sub-counties. Each of the counties had five (5) sub counties as described in the county profiles in Appendix 2. In Bomet, one (1) sub county was selected while two (2) sub counties were selected in Nyamira. The allocation was based on the number of tea growers in the two counties (Table A.1 in Appendix 1). The selected sub–counties were; Konoin in Bomet and Borabu and Masaba North in Nyamira. Konoin which was randomly selected is among the two tea growing sub counties in Bomet, while in Nyamira, Borabu and Masaba North were selected to represent areas with relatively larger and smaller farms respectively.

The third stage involved selection of four (4) administrative divisions; two (2) from Konoin in Bomet and one (1) each from Borabu and Masaba North in Nyamira. The selection of the divisions at this stage was random based on the list of administrative divisions in the respective sub counties. The selected divisions in each sub county were; Boito and Kimulot in Konoin, Nyansiongo in Borabu and Gesima in Masaba North.

The final stage involved the random selection of tea farming households. The sample size of 525 households was proportionately distributed among the selected divisions based on the population of tea farming households. The number of households assigned to each division was; 86 in Kimulot, 105 in Boito, 120 in Nyansiongo and 214 in Gesima. These households were randomly selected and interviewed.

3.7 Data Types and Sources

This study used primary data which was obtained through a cross-sectional survey of the smallholder tea farmers in Bomet and Nyamira counties in Kenya. The survey was conducted between December 2015 and March 2016.

3.8 Data Collection Instruments

During the survey, interviews with the respondents were undertaken using a structured questionnaire. To ensure suitability of the questionnaire, a pilot study was undertaken involving 15 and 20 households in Bomet and Nyamira County respectively. The pilot study was also intended to ensure that the survey enumerators understood all the questions in the questionnaire, how to ask

questions and the expected information, and how to handle problems that would arise during the interviews. From the Pilot study it was established that the questionnaire was adequate for collection of the data which was required to address study's objectives.

3.9 Data Cleaning, Coding and Presentation

After the field interviews, the completed questionnaires were scrutinized to assure that the collected data was devoid of errors, omissions or inconsistencies. This was done in the field in order to allow corrections and verification with the respondents. After all the questionnaires were completed, coding of the qualitative responses was done in order to facilitate analysis. The collected information was then applied in analysis using DEAP and STATA statistical packages.

3.10 Data Analysis

The study used linear programming and regression analysis techniques to meet the objectives of the study. Before analysis of the objectives, descriptive statistics were computed on the various variables applied in the study's models, in order to develop understanding on the variability in the variables. The first objective which sought to estimate the TE of tea farms in the selected counties was achieved using the linear programming DEAP statistical software for DEA. The DEA model (specified in section 3.4.1) was estimated and solved 525 times to obtain TE scores for all the farms in the sample. The estimation utilised the VRS assumptions which allowed for the decomposition of TE into pure TE and scale efficiency (SE). The scores were then averaged to obtain the level of TE of tea farms in the study area. The model was estimated in stages to allow for characterization of returns to scale

and analysis of slacks. After estimation, the average TE scores for Bomet and Nyamira counties were also computed and compared to obtain the differences in TE performance in the two counties.

The second objective sought to determine the effect of farm size on TE. The fractional model specified in 3.32 was estimated using quasi-maximum likelihood procedure (QML) based on the Bernoulli log-likelihood function which is efficient and consistent among the class of estimators containing all the exponential family based QML estimators (Papke and Wooldridge, 1996).

A part from determining the effects of farm size on TE, the model was also used to ascertain the influence of the other factors on efficiency. The robustness and validity of the FRM model was ascertained using alternative model specifications.

The third objective sought to analyse the effect of marketing arrangements on TE. To address the objective; the ESR model specified in equation 3.33 to and 3.35 was estimated using the full information maximum likelihood (FIML) estimation. The effect of green leaf marketing arrangements on TE was derived as the difference in TE for households participating in the alternative tea market channels and the TE the same households would attained had they not participated. The difference was obtained using the simulation specified in equation 3.15. Robustness of the ESR results was checked using the propensity score matching (PSM) model.

3.11 Diagnostic Tests

Before adoption and interpretation of the estimated models various diagnostic tests were conducted in order to ensure that obtained estimates were unbiased and consistent. The diagnostic tests that were conducted included; the variance inflation factor (VIF) test to confirm absence of severe multicollinearity, Link test for specification errors, the likelihood ratio (LR) test for good ness of fit, Wald test for joint significance of coefficients and the Lagrangian multiplier (LM)test for heteroskedasticity.

CHAPTER FOUR

EMPIRICAL FINDINGS AND DISCUSSIONS

4.1 Introduction

This chapter presents the study's findings and discussions. The chapter includes a presentation of the descriptive statistics for the variables that were included in the estimated models, the results of diagnostic tests on the models and the empirical results and their discussions.

4.2 Descriptive Statistics on Data

The dataset used in this study was based on household level data collected between December 2015 and March 2016 from a sample of 525 tea farming households in Bomet and Nyamira counties in Kenya. To describe the basic features of the data, descriptive statistics in terms of the means, measures of dispersion and frequencies were used as presented in tables 4.1 and 4.2.

Table 4.1 Summary Statistics for the Continous Variables

Variable	Total sample (N= 525)				Bomet (N=194)	Nyamira (N=331)	t test for difference in means
	Mean	Std dev	Min	Max	Mean	Mean	t score
Farm size (acres)	1.3	1.1	0.1	7	1.6	1.2	3.8
Fertilizer per farm (bags)	4.6	3.8	0.0	30	5.8	3.9	5.8
Labor per farm (Mandays)	163.9	138.5	3.0	1032	136.2	180.1	3.5
Labor structure (%)	58.2	45.5	0.0	100	62.3	55.7	1.6
Age of farm (Years)	27.0	14.9	3.0	60	21.8	30.1	6.4
Output per farm (Kgs)	3208.3	3019.8	100.0	14400	5035.9	1820.0	14.0
Yields per farm(kgs/acre)	2745.9	2067.6	100.0	10000	3907.9	2064.9	10.9
Household size (No)	6.3	2.8	1.0	20	6.0	6.5	1.8
Age of farmer (years)	49.2	14. 5	19.0	89	45.3	51.5	4.8
Household expenditure	237507.	160140.3	30920	1284000	240948.0	235516.0	0.4
Per capita expenditure	42658.0	32682.0	3992.9	370400	45602.3	40954.7	1.6
Household Assets (Kes)	107245.8	98976.6	1800.0	585200	117674.6	101379.6	1.8
Food sufficiency (Months)	9.5	4.3	0.0	12	10.3	8.1	5.8
Distance to market (kms)	2.90	2.73	0.1	20	2.84	2.93	0.39

The differences in the means are significant for values of t above 1.8
Source: Author's computation

Table 4.1 shows that within the sample, the operated tea farms were relatively small ranging between 0.1 and 7 acres with a mean of 1.3 acres. This observation was expected in the context of on-going subdivision of farms that is prevalent in the areas covered by the study (Republic of Kenya, 2014b). The average farm size under tea in the two counties was 1.6 and 1.2 acres for Bomet and Nyamira County, respectively. The difference in the mean size of tea farms between the two counties was tested using the t statistic and found to be statistically significant. This shows that smallholder tea farmers in Bomet operated relatively larger farms than their counterparts in Nyamira; an observation which is consistent with the demographic structure in the two counties (Republic of Kenya, 2009).

Apart from land, fertilizer and labour are the other key inputs used in tea production. Among the sampled farmers, the average amount of fertilizer applied was 4.6 bags of 50 kilogram (kgs) fertilizer per farm with a standard deviation of 3.8 bags. This level of fertilizer is below the recommended annual rate of 5 bags per acre considering that the average farm size was 1.3 acres (TRFK, 2002). The average annual labour utilized per farm in the overall sample was 163.9 man-days compared to a mean of 136.2 man-days for Bomet County and 180.1 man-days for Nyamira County. The results show that while the smallholder farmers in Bomet County used more fertilizer, those in Nyamira had higher levels of labour use. Although causality cannot be ascribed, the use of more labour in Nyamira is consistent with the expectation that more densely populated areas are likely to experience less labour supply bottlenecks. In terms of the structure of labour, the results show that on average 58.2 percent of the labour used in tea production was

supplied by the household while the remaining 41.8 percent was hired. The difference in the average proportion of household labour between the two counties was statistically insignificant, which indicates that the labour structure in the two counties was relatively similar.

At the farm level, the age of the tea farm is an important farm characteristic since aging tea plantations are associated with decline in tea productivity (Do and Le, 2000; Kamau, 2008). The average age of tea bushes in the sample was 27 years. Comparatively, the results revealed that the difference in the mean age of tea bushes in the two counties was statistically significant. The mean age was 21.8 years for Bomet and 31.1 years in Nyamira County. This implies that tea bushes in Bomet County were relatively younger and therefore expected to be more productive than those in Nyamira County.

With regard to the tea output, the mean annual output per farm was 3208.3 kgs of tea with a standard deviation of 3019.8 kgs and a range of 100 kgs to 14400 kgs. The mean annual tea output per farm for Bomet was 5035.9 kgs compared to 1820 kgs for Nyamira County. The mean annual yield for the overall sample was 2745.9 kgs per acre compared to 3907.9 kgs and 2064.9 kgs per acre for Bomet and Nyamira County respectively. This shows that farmers in Bomet County had higher annual yields per acre than their counterparts in Nyamira County. The superior performance in yields by Bomet County would be attributed to the application of higher fertilizer rates, presence of younger and more productive tea bushes and differential utilization of agricultural credit (Table 4.1).

Regarding the household demographic profile, the results showed that the average household had six members with a minimum and maximum of 1 and 20 members respectively. The mean age of household head was 49.2 years and ranged from 19 to 89 years. This implies that tea farmers are relatively aged; an observation that is consistent with a preference by younger people to engage in non-farm enterprises (Ogundari *et. al.,* 2006; Kiprono, 2013). The average age of the household head in Bomet was 45.3 years compared to 51.5 years in Nyamira. The difference in the mean age of the farmers between Bomet and Nyamira was statistically significant implying therefore that the farmers in Bomet were relatively younger than those in Nyamira County.

The average annual household per capita expenditure was Ksh 42,658, while the average value of household assets was Ksh 107,245.80. On food sufficiency, the results revealed that the average period over which the household relied on the market to access food staples was 9.46 months. The mean value for the variable was 10.3 months for Bomet and 8.1 months for Nyamira County, implying that households in Nyamira had higher levels of food sufficiency than the households in Bomet. The average distance from the farm to the nearest market was 2.9 kilometres with a standard deviation of 2.7 kilometres.

The other variables which the study considered were categorical in nature and their summaries are presented in table 4.2. The variables include gender and residence of the household head, the highest level of education attained by the household head, the variety of tea planted, soil fertility, access to extension services and the green leaf marketing channel.

Table 4.2 Summary Statistics for the Categorical Variables

Variable	Measurement	Combined (n=525)		Bomet (n = 194)		Nyamira (n = 331)		Z test
		Freq	%	Freq	%	Freq	%	Z score
Gender of household head	Male	442	84.2	168	86.6	274	82.8	1.16
	Female	83	15.8	26	13.4	57	17.2	1.16
Residence of the household head	Yes	469	89.3	188	96.9	281	84.9	4.30
	No	56	10.7	6	3.1	50	15.1	4.30
Education level of household head	Primary	238	45.3	91	46.9	147	44.4	0.55
	Secondary	211	40.2	70	36.1	141	42.6	1.47
	Tertiary	50	9.6	24	12.4	26	7.9	1.70
	University	26	5.0	9	4.6	17	5.1	0.25
Credit utilisation	Yes	363	69.1	147	75.8	216	65.3	2.51
	No	162	30.9	47	24.2	115	34.7	2.51
Adopted clonal tea varieties	Yes	315	60.0	133	68.6	182	55.0	3.06
	No	210	40.0	61	31.4	149	45.0	3.06
Soil fertility	Suitable	371	70.7	149	76.8	222	67.1	2.70
	Not suitable	154	29.3	45	23.2	109	32.9	2.70
Extension access (T&V)	Yes	430	81.9	141	72.7	289	87.3	4.20
	No	95	18.1	53	27.3	42	12.7	4.20
Extension access (FFS)	Yes	277	52.8	107	55.2	170	51.4	0.84
	No	248	47.2	87	44.8	161	48.6	0.84
Greenleaf market channel	KTDA	334	63.6	109	56.2	225	68.0	2.71
	ATMC	191	36.4	85	43.8	106	32.0	2.71

The differences in the means are significant for values of t above 1.8
T&V means train and visit model of extension
Z test if fortesting difference in % between Bomet and Nyamira
Source: Author's computation

Table 4.2 shows that 84.2 percent of the sampled households were headed by males while the remaining 15.8 percent were female headed. This result is within the range reported in other Kenyan household based studies and is consistent with the expectation that most farm households in developing countries are male headed (Republic of Kenya, 2005-06; Republic of Kenya, 2014b). A part from gender, the residence of the household head is the other household characteristic which the study considered. This variable was included as an indicator of the availability of the head of the household on full time basis to make the relevant

farm decisions. The results show that 89 percent of the household heads in the sample resided within the homestead on a full time basis and were therefore available at the farm to make the relevant decisions in tea farming. Only 10.7 percent of the sample did not reside within the homestead on a regular basis. Comparatively the proportion of household heads that resided within the homestead was 96.9 percent in Bomet compared to 84.9 percent for Nyamira County.

On education, the results show that 45.3 percent of household heads in the sample had attained primary level education, 40.2 percent had secondary level education and 9.6 percent had attained college education. Only 5.0 percent of the sampled household heads in the sample had university education. The results imply that majority of the sample farmers had not attained higher education in terms of college or university education. In terms of use of credit, the results revealed that 69.1 percent of the interviewed households had utilized agricultural credit while the remaining 30.9 percent had not. The proportion of households that had used credit in Bomet County was 75.8 percent compared to 65.3 percent in Nyamira County. The difference in the levels of utilization would mean that farmers in Nyamira County were either more constrained in accessing credit or were more averse to its use.

At the farm level, the planted tea variety has significant influence on yields. The study considered whether the household had planted the improved clonal varieties or the non-improved seedling tea varieties. The results show that 60.0 percent of the sampled households had planted improved tea varieties while the remaining

40.0 percent had not. The proportion of households that had adopted improved clonal varieties in Bomet County was 68.6 compared to 55.0 percent in Nyamira County. The higher rate of adoption observed in Bomet County may to some extent explain why the county had achieved better performance in terms of tea yields.

The higher adoption rate of tea clones in Bomet County could be attributed to county's proximity to the Tea Research Institute (TRI). The institute is responsible for the release of tea clones to the farmers and is located in Kericho. The proximity of Kericho to Bomet may have made the access of the improved clonal materials to be easier among farmers in Bomet than in Nyamira County. The other reason for the observed differences in adoption is that tea farms in Nyamira are relatively older than farms in Bomet. This implies that most of the tea bushes in Nyamira could have been established prior to 1978 when the tea cloning technology was introduced in Kenya. It is worth to note that tea is a perennial crop and its replanting involves significant costs.

Tea may be grown in soils of diverse origins. In Kenya, the crop performs best in soils of volcanic origin that are deep (1.8-2 metres), well drained and are brownish red or dark red in colour (TRFK, 2002). On soil fertility, the study considered whether the soil fertility of the tea farm was suitable for tea production based on self-assessment by the farmers. The indicators used in the fertility assessment were those that the farmers could see or feel including soil depth, colour, drainage, slope and weed abundance. The approach was based on literature indicating that although farmers' knowledge is often ignored in soil studies, farmers do have good

ability to perceive differences in the level of fertility between and within fields on their farms (Desbiez *et. al.,* 2005; Maro *et. al.,* 2013).

The results show that about 70.7 percent of the respondents considered that the soil fertility in their farms was suitable for tea farming. The remaining 29.3 percent of the respondents reported that the soil fertility in their farms was not suitable for tea farming. Comparatively, the proportion of respondents who reported that the soil fertility in their farms was suitable for tea farming was 76.8 and 67.1 percent in Bomet and Nyamira Counties, respectively. The differences in soil quality between the two counties might be the other factor that would explain the observed yield differences in the two counties.

Regarding the access to extension services, the results showed that 81.9 percent of sample households had access to the Train and Visit (T&V) extension system. The proportion of the sample households who had access to the T &V extension in Bomet County was 73 percent compared to 87 percent in Nyamira County. This means that farmers in the two counties had relatively good levels of access to the T&V based extension. In contrast, the results show that much lower levels of access were observed under the Farmer Field School (FFS) based extension system. The results show that the proportion of farmers who had access to the FFS extension systems was 52.8 percent.

With regard to the marketing of green leaf, the study considered whether the household exclusively sold its green leaf to a KTDA factory or used alternative market channels (ATMCs). The ATMCs considered in the study included the sale

of green leaf through middlemen or the multinational and private tea factories. The results show that the KTDA channel was used by majority of the households and accounted for nearly 63.6 percent of the sample. The proportion of households that used the alternative markets channels in the sample was 36.4 percent in the total sample, 43.8 percent in Bomet and 32.0 percent in Nyamira County. These results imply that the ATMC which is an outcome of the liberalization of the tea sector has since grown and currently commands a sizable share of the green leaf market.

Overall the results of the descriptive statistics, illustrate that the data used in the study was consistent with the broad body of smallholder data in Kenya. The data set was therefore adopted for analysis of the study's objectives.

4.3 Technical Efficiency of Smallholder Tea Production

The first objective of the study sought to estimate the technical efficiency (TE) of smallholder tea farms in the study area. This was achieved using the non-parametric data envelopment analysis (DEA) model which is less prone to misspecification in the analysis of TE, and also provides other additional information on scale efficiency (SE), slacks and returns to scale (Ji and Lee, 2010; Zhu, 2014). The results of the model in terms of the level and distribution of TE are discussed in section 4.3.1 while those on returns to scale and input slacks and are presented in section 4.3.2.

4.3.1 The Level and Distribution of DEA Scores

In order to estimate the level of technical efficiency (TE), the data collected from the survey was analysed using the DEA methodology described in chapter three. The TE scores were calculated under the assumptions of variable returns to scale (VRS). The results of the TE scores for each of the sample households are presented in table A.3 in the Appendix. The results show that the TE scores had a wide distribution ranging between 0.07 and 1.0. This observation is consistent with Kiprono (2013) who found that the economic efficiency of smallholder tea farmers in Kenya ranged from 0.01 to 0.74. The results are also similar to Nguyen-van and To-the (2014) who found that the distribution of the TE scores varied from 0.05 to 0.95 among smallholder tea farmers in northern Vietnam. The results are however inconsistent with Hong and Yabe (2015) who found that the TE of Vietnamese smallholder tea farmers ranged from 0.62 to 0.97. The frequency distribution of the TE scores across the sample is shown table 4.3.

Table 4.3: Frequency Distribution of The TE scores

Distribution of TE scores	Combined sample			Bomet			Nyamira		
	N	Percent	Cumulative (Cum) %	N	Percent	Cum %	N	Percent	Cum %
0- 0.10	4	0.8	0.8	0	0.0	0.0	4	1.2	1.2
0.11-0.20	62	11.8	12.6	10	5.2	5.2	52	15.7	16.9
0.21-0.30	93	17.7	30.3	19	9.8	15	74	22.4	39.3
0.31-0.40	91	17.3	47.6	27	13.9	28.9	64	19.3	58.6
0.41-0.50	98	18.7	66.3	39	20.1	49	59	17.8	76.4
0.51-0.60	50	9.5	75.8	28	14.4	63.4	22	6.6	83.0
0.61-0.70	42	8.0	83.8	25	12.9	76.3	17	5.1	88.1
0.71-0.80	29	5.5	89.3	13	6.7	83.0	16	4.8	92.9
0.81-0.90	12	2.3	91.6	8	4.1	87.1	4	1.2	94.1
0.91-1.00	44	8.4	100.0	25	12.9	100.0	19	5.9	100.0
Total	**525**	**100.00**	**100.00**	**194**	**100.0**	**100.0**	**331**	**100.0**	**100.0**

Source: Author's computation

From the results in table 4.3, it is apparent that smallholder tea farmers in the study area were comparatively less efficient with majority having relatively low TE scores. The proportion of the sampled farmers whose score was below 0.5 was 66.3 percent. Only 8.4 percent of the sampled households had obtained TE scores of more than 0.9.

A notable observation from the distribution in table 4.3 is that the farmers from Nyamira County were proportionately less efficient than their counterparts from Bomet County. For instance, the proportion of farmers whose TE scores were above 0.5 was 51 percent for Bomet compared to 24 percent in Nyamira County.

Under VRS assumptions, TE can be decomposed into two mutually exclusive and non-additive components; Pure TE and scale efficiency (SE). The first component reflects the managerial performance of the farm in organizing inputs in the production process. The latter component expresses whether a farm is operating on an optimal scale and provides a measure of the farm's ability to choose the optimal size. The decomposition is useful in providing insights about the sources of inefficiency since inappropriate size (too large or too small) may also be a cause of inefficiency in production (Kumar and Gulati, 2008). Table 4.4 shows the comparative analysis of the average TE and SE scores for the smallholder farmers in the two counties.

Table 4.4: The Mean TE scores for Bomet and Nyamira County

	Bomet (n =194)	Nyamira (n=331)	Difference		sample (n=525)	
	Mean	Mean	Difference	t statistic	Mean	Std Dev
Mean TE	0.55	0.41	0.14***	7.03	0.46	0.24
Mean SE	0.82	0.59	0.23 ***	9.38	0.67	0.28

Source: Author's computation Legend:

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)

As shown in table 4.4, the mean TE score obtained in the overall sample was 0.46 which implies that farmers in the study area are less successful in obtaining maximum output given the existing resources. The result means that by adopting the production practices of the best practice farmers, smallholder tea farmers can achieve their current tea production levels using only 46 percent of their existing resources. The average TE score estimated in this study was higher compared to Kiprono (2013) and Nguyen-van and To-the (2014) who found an average TE score of 0.15 and 0.32 among smallholder tea farmers in Kenya and northern Vietnam respectively. However the score is lower compared to Baten *et. al.* (2010), and Basnayake and Gunaratne (2002), who found the average TE score among smallholder tea farmers in Bangladesh and Sri Lanka to be 0.59 and 0.64, respectively.

It can further be observed that differences in the means of the TE and scale efficiency (SE) scores between Bomet and Nyamira were statistically significant at one percent level. This implies that smallholder tea farmers from Bomet performed significantly better than their counterparts from Nyamira County in terms of both the TE and SE measures. The better performance in TE by smallholder tea farmers in Bomet County would be attributed to the fact that the county had better

adoption levels of clonal tea varieties and relatively younger tea bushes. The other probable factors would be the more intensive use of fertilizer and better access to credit and FFS extension. The differences in TE between two counties imply that although tea growing areas appear to have relative similarities in agro-ecological conditions, there are performance disparities which need consideration when formulating policies in the tea sector.

From the SE scores, it is evident that smallholder tea production is not scale neutral and the contribution of scale is an important source of efficiency variation in tea production. The source of scale inefficiency was analysed in terms of whether a farm was characterized by increasing or decreasing return to scale. The results of the RTS characterization are discussed in section 4.3.2.

4.3.2 Analysis of Returns to Scale and Level of Inputs Slacks

One of the key advantages of DEA is that it allows for the identification and analysis of returns to scale. Understanding of returns to scale is important in indicating whether any efficiency gains can be obtained by adjusting the size or scale of operation of a farm (Tolga *et. al.,* 2009). For instance, farms experiencing increasing returns to scale can benefit by becoming larger or similarly farms at the optimal scale can suffer efficiency losses if the scale of production is adjusted. Table 4.5 shows the distribution of the DEA model results on return to scale among the sampled tea farms.

Table 4.5: The Distribution of Returns to Scale from the DEA Model

Returns to scale	Frequency	Percent	Cumulative
Decreasing (DRS)	44	8.38	8.38
Increasing (IRS)	470	89.52	97.9
Optimal scale (OS)	11	2.1	100.0
Total	**525**	**100**	**100.0**

Source: Author's Computation

The results revealed that the smallholder tea farms in the study area showed different returns to scale characteristics with 8.3 percent of them exhibiting decreasing returns to scale. Further, the results show that majority of the farms (89.5 percent) were characterized by increasing returns to scale (IRS), suggesting therefore that there are scale advantages in tea production which can be harnessed by increasing the scale of operation of the smallholder farms. Within the sample, only 2.1 percent of the farms were operating on the optimal scale. The findings point to the significance of farm size issues in tea farming.

In order to allow for a more comprehensive measure of efficiency, the analysis of TE in the study was extended to capture the additional inefficiencies represented by non-zero slacks. This was necessary since in DEA, non-zero slacks can represent a substantial amount of inefficiency such that a firm is considered efficient when it's DEA score is equal to one and all its slacks are equal to zero (Seiford and Zhu 1999; Morita *et. al.,* 2005). The non-zero slacks therefore relate to the input excesses and show the extent to which further reductions in inputs could be achieved beyond the levels implied by the results of the TE scores (Coelli *et. al.,* 1996). While the level of TE represents the amount by which each input may be reduced without changing the proportions in which the inputs were used,

any further reductions evidenced by input slacks must change the input proportions (Banker *et. al.,* 2004). The analysis of input slacks was achieved by implementing the second stage of the DEA procedure specified in equations 3.24 to 3.31. The results of the analysis of the input slacks in terms of the mean and frequency distribution are shown in table 4.6.

Table 4.6: Analysis of Input Slacks from the DEA Model

	Frequency of Slack Observations in the Sample			
Input	**Zero (0)**	**Non-zero**	**%**	**Mean**
Farm size	507	18	3.4	0.032
Fertilizer	512	13	2.5	0.065
Labour	508	17	3.2	0.125

Source: Author's Computation

The results reveal that only a small proportion of the sample farms (less than 4%) had non-zero input slacks in all the three (3) inputs. In addition the mean values were equally low, which led to the conclusion that slacks were not very important in characterizing efficiency in smallholder tea production in the study areas. This is consistent with the observation that slack issues tends to disappear as the number of DMUs increases since the piecewise linear frontier becomes smoother and therefore minimizes the likelihood of it running parallel to the axes (Coelli *et. al.,* 1996).

Overall, the results of TE analysis show that smallholder tea farmers in the study areas were technically inefficient and that there exists considerable potential to improve the wellbeing of tea farmers through improvement in TE. The results further reveal that smallholder tea farmers in the study area are less successful in

employing the production technology of the best practice farmers than are the farmers in other countries such as Sri Lanka and Bangladesh. In addition, the results indicate existence of location specific heterogeneities that influence the attainment of efficiency in smallholder tea production. Finally the results on efficiency analysis indicate that majority of the smallholder tea farms in the study area were not operating at or near to their optimal scales.

4.4 Effects of Farm Size on Technical Efficiency

To examine the effects of farm size on technical efficiency (TE), the TE scores obtained for each household were regressed against farm size and a set of other explanatory variables which included the age of the tea farm, distance to the nearest market, education and access to credit. The other covariates included access to extension, a county dummy to cover the regional fixed effects and green leaf marketing channel. These other variables were included in the model so that apart from understanding the effect of farm size on TE, their effect could also be ascertained.

The model specified in section 3.3 in chapter three was estimated using the fractional regression (FR) model proposed by Papke and Wooldridge (1996) to model bounded and fractional dependent variables (Wooldridge 2002; McDonald 2009; Ramalho *et. al.,* 2010). The choice of the model was justified since it is based on a functional form that imposes the restriction that the conditional mean of the dependent variable is bounded within the unity interval which is consistent with analysis of DEA scores. Using the alternative standard linear model would have produced biased parameter estimates because of the possibility that the

predicted TE values can lie outside the unit interval making them incompatible with the bounded nature of the TE scores (Muniu *et. al.,* 2015). The Tobit model which has been applied in some studies would also have been inconsistent since its domain applies to situations when the dependent variable has been censored or an outcome of an optimization problem for which there are corner solutions (Cameron and Trivedi, 2005; McDonald 2009; Ramalho *et. al.,* 2010) .

Before estimating the model, a correlation analysis of the explanatory variables was done to check for the presence of multi-collinearity among the explanatory variables. The results of the correlation analysis are presented in table A.4 in Appendix. The results revealed that all the explanatory variables selected for inclusion in the FR Model were not strongly correlated. This therefore led to the conclusion that there would not be any severe multi-collinearity in the model under estimation. The model was therefore estimated and marginal effects obtained. The output of the estimated FR model is shown in the appendix in Table A.5.

4.4.1 Results of Diagnostic Tests

After estimation, the model was tested for specification and overall significance, heteroskedasticity, multicollinearity and robustness. The tests which were conducted include the link test to detect specification errors in the model, the Wald test to check for overall goodness of fit and whether the coefficients of the variables included in the model were jointly statistically significant and the Lagrangian multiplier (LM) test for heteroskedasticity. The results of the diagnostic tests are presented in table 4.7.

Table 4.7: Results of Diagnostic Tests

Test	Ratio/value	P value
Link test for model specification:		
Hat coefficient	1.569	0.000
Hat^2 coefficient	-0.244	0.399
Wald test : Chi ^{2} (14 df)	200.49	0.000
LM test **:** Chi 2 (14 df)	8.63	0.656

Source: Author's estimation

Table 4.7 shows that the coefficient of the linear predicted value of the model (hat) for the link test was significant (p=0.000) while coefficient of the square of the linear predicted value (hat^2) was insignificant (p=0.399). The test was based on the null hypothesis that the model was specified correctly. Based on the p values of the coefficients, the test failed to reject the assumption that the model was specified correctly and concluded the model did not suffer from misspecification. The intuition is that the square of the linear predicted value (hat2) would have had predictive power if the model was incorrectly specified. The Wald test applied in the study was based on the null hypothesis that all the coefficients of the model were simultaneously equal to zero. The chi-square statistic for the test had a probability value of 0.000 which led to the conclusion that the coefficients of the variables included in the model were jointly statistically significant. To test for heteroskedasticity, the study used the LM statistic to test the null hypothesis that the coefficients of an auxiliary regression involving residuals of the FR model were jointly equal to zero. The results of the test show that the test statistic was insignificant which confirmed absence of heteroskedasticity.

In addition to the specification tests and heteroskedasticity tests, the variance inflation factor (VIF) test was implemented to confirm absence of severe

multicollinearity in the estimated model. The VIF test measures how much the variance of estimated regression coefficients is inflated compared to when the explanatory variables are not linearly correlated (Verbeek, 2012). The results of the test (Table A.6 in the Appendix) show that the mean VIF was 2.99 and the values of the individual explanatory variables were not large (less than 10). This implied that null hypothesis of no severe multicollinearity could not be rejected and therefore attested to the absence of severe multicollinearity in the estimated model.

Before adopting the model to address the study's objective, checks on the robustness of the model and validity of its results were also undertaken. The checks were based on two different approaches. The first approach was implemented using a bootstrapping procedure for the model's standard errors. The bootstrapping procedure allows resampling by taking of samples over and over again from the same sample data in order to provide an indication of how accurate the original sample estimates were about the entire population. The procedure simulates repeated sampling of the population based on the principle that sampling with replacement behaves on the original sample the same way the original sample behaves on the population (Hall, 1994). The results of the re-estimated model based on 50 replications are presented in the appendix in Table A.7. The results show that the effect of farm size on TE had the same sign and significance as in the FR model. Similarly, the coefficients of the other variables were similar in sign and significance, which therefore confirmed that the results from the DEA and estimated FR models were robust.

The second test was to check whether inclusion of farm size as an input in the DEA estimation had caused biases in the estimated relationship between farm size and TE. This was achieved by re-estimating the FR model using DEA scores that were recomputed after excluding farm size from the input vector. The approach was implemented following Helfand and Levine (2004). The results of the recomputed model were consistent with the FR model in terms of the sign and significance of the coefficients for farm size and its quadratic term (Table A.8 in the Appendix). Except for square of age of the farm, education (secondary) and labour structure, all the coefficients of the other variables in the recomputed model maintained their sign and significance, therefore reaffirming the robustness of the model.

From the results of the alternative specification, it was concluded that the FR model which used to estimate TE and its relationship with farm size was robust and was therefore adopted for addressing the second objective of the study.

4.4.2 The Marginal Effects of Farm Size and other variables on TE

The marginal effects of farm size and other various variables on TE are presented in table 4.8.

Table 4.8: The Marginal Effects from the FR Model

Dependent Variable: TE Score				
Variable	**Coefficient**	**Std. error**	**Z value**	**P value**
Farm size	-0.165***	0.031	-5.27	0.000
Square of farm size	0.021***	0.006	3.35	0.001
County dummy	0.157***	0.021	7.37	0.000
Variety of tea	0.011	0.019	0.58	0.561
Age of farm	-0.006**	0.003	-2.23	0.026
square of age of farm	0.000	0.000	1.76	0.078
Gender of farmer	-0.015	0.028	-0.55	0.585
Education (primary)	0.080**	0.035	2.29	0.022
Education (secondary)	0.074**	0.034	2.17	0.030
Education (college)	0.077*	0.040	1.91	0.057
Extension (FFS)	0.053***	0.019	2.79	0.005
Distance to market	-0.002	0.004	-0.53	0.598
Labour structure	0.046*	0.024	1.91	0.055
Market channel	0.074***	0.021	-3.58	0.000

Legend:
- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- Marginal effects for dummy variables refer to the discrete change from 0 to **1**
- The default education level is university
- Source: Author's computation

The results in table 4.8 show that the coefficient for farm size was negative and statistically significant at one percent level. The coefficient of the quadratic term which was introduced to examine whether the negative effects of farm size on TE would persist at all levels of farm size was positive and statistically significant at standard levels. The results show that the effect of farm size on TE in smallholder tea production is non-linear with TE first falling and then rising with increase in farm size. Specifically the results imply that while TE initially decreases with rise in farm size, there is a unique turning point (defined by farm size) beyond which TE rises with increase in farm size.

The critical level of farm size beyond which the sign of the marginal effect is reversed was obtained by taking partial derivative of the TE equation with respect to farm size and then equating the result to zero. The critical level of farm size was then solved as follows;

$$\frac{\partial TE}{\partial Farmsize} = -0.165 + 2(0.021 farmsize) = 0$$

*Critical Farm-size** = 3.93 acres

The calculated level of farm size (3.93 acres) represents the scale of operation where the advantages of small farms would be outweighed by the economies of scale of the larger farms. Specifically, the result means that for farms whose acreage was below the critical farm-size, TE would diminish with increase in farm size. Similarly, TE would increase with an increase in farm size for farms whose acreage was above the calculated critical farm size. This result is inconsistent with Chirwa and Kydd, (2006) who found an inverse and continuously linear relationship between productivity and farm size among smallholder tea farmers in Malawi. The result also contradicts Maity (2012) who found that TE increased with the size of tea gardens in West Bengal in India.

As shown in table 4.8, the value of the farm size coefficient was -0.165. This implied that increasing the area under tea by one acre for farms whose acreage was less than 3.93 acres, would lead to a reduction in the TE score by 0.165. This negative effect could be attributed to the more intensive use of labour on smaller farms than the larger farms. This explanation was considered plausible since tea

farming is labour intensive and has a calendar of activities that run throughout the year (Chirwa and Kydd, 2006). The validity of the explanation was examined by correlating farm size with the amount of labour used in tea production per unit area. The analysis was done to test whether indeed smaller farms were using more labour per unit area than the bigger farms. The analysis excluded farms whose acreage was above 3.93 acres since the hypothesized association would only apply where the tea acreage was below the critical level of farm size. The results of the analysis are shown in the correlation matrix in table 4.9.

Table 4.9: Correlation Analysis: Farm Size and Labour Use in Tea Production

Variable	**Farm size (acres)**	**Labour (man-days per acre)**
Farm size (acres)	1	-0.4306*
Labour (Man days per acre)	-0.4306*	1

Denotes significance at 5% **Source: Author's computation**

Table 4.9 shows that there existed a negative association between farm size and the quantities of labour used in tea production per unit area. The result therefore affirmed the assertion that smaller farms had higher labour to land ratios than the bigger farms. The intensity of labour in the smaller farms would be attributed to the existence of imperfections in labour markets which make smaller farms to face cheaper labour costs than the large farms. The reason is that bigger farms rely more on hired labour which is normally more expensive than family labour due to additional transaction costs of search and screenings (Carter and Wiebe, 1990; Kiani 2008). In addition, the higher transaction costs may be associated with the additional monitoring costs since work effort in a tea farm may not be completely observable, verifiable and enforceable (moral hazard).

Apart from labour use, the other factor that would explain the inverse relationship between farm size and TE is the heterogeneity in land quality. The variable would explain the inverse relationship if land quality factors such as soil fertility are inversely correlated with farm size (Bhalla and Roy, 1988). The study adopted the t test for comparison of means to test whether land quality based on soil fertility was inversely correlated with farm size. The results of the analysis are presented in Table A.9 in the Appendix. The results show that the tea plots that were reported to have good soil fertility were relatively bigger than those whose soils were reported to be unsuitable for tea farming. This showed that land quality heterogeneity may not be an important factor in explaining the negative effect of farm size on TE in the study area.

The finding that the effect of farm size on TE is non-linear with TE first falling and then rising with increase in farm size contradicts Hong and Yabe (2015) who found that farm size did not have any important effect on TE in tea production in Vietnam. The finding is also at variance with the inverse productivity relationship that is reported among various agricultural enterprises in various countries (Rios and Shively, 2005; Masterson, 2007; Githinji *et. al.,* 2011; Chand *et. al.,* 2011). The finding is however consistent with the tea farm consolidation and enterprise diversification programs envisaged in the proposed National Tea Policy (Republic of Kenya, 2014a). From the findings of this study it can be concluded that while sub-division of tea farms may not per se have a constraining influence on TE, efficiency improvements can be achieved through consolidation of smallholder tea farms.

Overall, the results of this study reveal that farm size has important effects on TE, with efficiency first declining and then rising with increase in farm size. Apart from determining the effect of farm size on TE, the FR specification was also designed to ascertain the influence of the other covariates, including a county dummy, variety of tea, age of tea farm, household type, highest level of education attained by the household head, extension services, distance to markets, labour structure and market channel.

As shown in table 4.8, the coefficient for the county dummy was positive and statistically significant at one percent level. The variable was included to capture the effects due to location specific regional heterogeneities. The results show that on average smallholder tea farmers in Bomet were technically more efficient than the smallholder tea farmers in Nyamira County. This result is in line with the finding in table 4.4 that show that smallholder tea farmers from Nyamira were technically less efficient than their counterparts from Bomet. In addition to the differentials in the access to services and age of the tea farms, an arrangement in Nyamira County that limits tea plucking to five days per week (as opposed six day in Bomet County) may also explain the performance disparities. This result therefore suggests existence of region specific heterogeneities which should not be ignored when formulating policies in the tea subsector. The observation is consistent with Kavoi *et. al.* (2001) and Kiprono (2013) who found that smallholder tea farmers from the west of the Rift Valley were less efficient compared to smallholder farmers from the east of the Rift Valley.

Although the adoption of high yielding crop cultivars is thought to be one of the major sources of productivity growth in agriculture, the regression results in table 4.8 show that the coefficient for the variety of tea was statistically insignificant at standard levels of significance. The result suggests that the variety of tea is not an important determinant of TE in smallholder tea production in the study area. This finding is consistent with Kiprono (2013) who found that the tea variety was not an important source of efficiency variation in tea production in Kenya. The finding may be due to the possibility that the gains from the improved varieties would have been countervailed by the fact that traditional seedling varieties are more adaptable to climatic variability and adverse growth conditions. This is due to the ability of the seedling varieties to develop a vertically descending tap root unlike the clonal varieties which have a tendency to develop a spreading root system within the fertile and upper layers of the soil (Wickramaratne, 1981).

The age of the tea farm was the other factor considered in the study. The economic significance of age derives from the fact that peak yields for tea are obtained between ages of 25 to 40 years after planting, followed by a decline to a level where the tea gardens may become moribund (Kamau, 2008). The coefficient for the variable was negative (-0.006) and statistically significant at 5 percent level, implying that an increase in the age of the tea farm by one year would lead to a reduction in TE by 0.006. This finding is consistent with experimental studies that have shown that younger tea plantations are generally associated with higher productivity in most tea growing regions in the world (Mwakha, 1989; Bore, 1996). The study's finding is also in conformity with tea industry's assertions that

aging tea gardens were causing stagnation in tea productivity (Republic of Kenya, 2014a). To reverse the decline in yields due to aging of tea bushes, other countries such as Sri Lanka, India and Malawi are implementing tea replanting programs for their smallholder farmers.

On the level of education attained by the household head, the results show that the coefficients for tertiary, secondary and primary were all statistically significant. The results imply that farm households whose head had lower than university education (primary, secondary and tertiary) had higher levels of TE than farm households headed by individuals with university education. These results seem to suggest that higher education (university education) has a negative effect on TE in smallholder tea production. This finding is inconsistent with the expectation that education should improve access to information and understanding of tea husbandry practices. A plausible explanation for the finding could be related to the differential allocation of time to tea farming activities, given that the more educated farmers have a higher opportunity cost of labour because they can earn higher returns outside farming (Nyangena, 2008; Yoshiko, 2010). It is therefore likely that farmers with university education were allocating less time to tea production activities than the less educated farmers.

The coefficient for gender of the household head was insignificant implying that the TE of farm households headed by females was not statistically different from the TE of the farm households headed by males. The finding is in contrast with Hong and Yabe (2015) who found that male headed households had higher levels of TE than the female headed households. The finding is however consistent with

Quisumbing (1996), whose survey on differences in TE between male and female farmers in agricultural production found that six in seven studies had insignificant dummies for gender of the household head. The finding is also consistent with Chirwa and Kydd, (2006) who found no statistical evidence of gender differences in tea productivity in Malawi.

With regard to extension services, table 4.8 shows that the coefficient for the FFS extension was positive and statistically significant at one percent level. This implies that the TE of tea farms household that had participated in the FFS program was higher than that of the non-participating farm households. The results are consistent with Onduru *et al.* (2012) who found that participation in FFS had a positive and significant influence on tea yields in selected tea growing regions in Kenya. This finding was expected since FFS is a new approach to extension which was adopted by the tea sector to address the weaknesses of the conventional T&V approaches (Mose *et. al.,* 2016). The attractiveness of FFS is associated with its use of participatory adult leaning approaches and emphasis on stronger linkages between research, extension and farmer experimentation (Friis-Hansen, 2004). The coefficient of access to T&V extension was insignificant in all the earlier specifications that were tried and was therefore dropped in the final model.

In the study, distance to markets was included as an indicator of market related transaction costs since an increase in the distance to markets can act as an economic disincentive to the farmers. However the estimated coefficient for distance to nearest market was not statistically significant. This might be explained by the recent improvements in road infrastructure in the country and the

emergence of motorbikes as a popular mode of transport in rural areas. The improved access to market as a result of these developments would have diminished the influence of distance to markets as a variable explaining TE in tea farming.

The coefficient of labour structure was positive and significant at 10 percent implying that an increase in the share of family labour applied in tea production is associated with an increase in TE. The finding is consistent with Ogada *et. al.* (2014) who found that households that utilized hired labour in food production in Kenya had lower TE scores than households that used family labour. The finding also affirms the earlier assertion that labour use is an important variable in tea production. Studies that have underscored the significance of labour constraints in tea farming include Owuor (2001) and Owuor *et. al*. (2005). The significance of labour in tea farming derives from the view that when labour access constraints are binding, farmers may fail to carry out the required tea husbandry practices at the optimum time. This is more prevalent during the rainy season when the tea yields are at their peak.

Regarding the marketing of green leaf, the coefficient for market channel was positive and statistically significant at one percent. The value of the coefficient was 0.074; which would imply that the TE of farm households that had participated in the alternative (non-KTDA) tea market channels was higher than the TE of farm households that had not participated. The coefficient however may not consistently estimate the causal effect of participation on TE, due to potential endogeneity since the decision to participate in alternative market channels is

voluntary and therefore subject to self-selection biases. In addition, a causal interpretation on the dummy would be misleading since the participants and non-participants would still exhibit differences in TE even in the absence of participation making a causal interpretation of estimated coefficient difficult (Verbeek, 2012). Therefore while the market channel coefficient in the FR model provides some insights on how ATMC participation affects TE; a more detailed analysis which accounts for unobserved farmer characteristics and self-selectivity in the decision to participate is provided in section 4.5.

4.5 Effect of Green Leaf Marketing on Technical Efficiency

The Kenya smallholder tea sub-sector was liberalized in the year 2000 with the objective of enhancing competitiveness and efficiency in the sub-sector. A major outcome of the reform was the advent of alternative tea markets channels (ATMC) for selling green leaf. The alternative markets operate in some parts of the tea growing areas in the country alongside the more dominant KTDA system and therefore allow farmers options on where to sell their tea. Where the alternative markets operate, tea from the smallholder farmers is sold through three ATMCs. In the first channel, some farmers especially those in the catchment areas neighbouring the large scale tea estates directly sell their green leaf to the factories owned by these estates. The second channel constitutes the sale of green leaf through the middlemen in spot markets along the roadside or at the farm gate. The third channel involve farmers who participate in ATMCs through intermediary institutions such cooperatives and farmer organizations.

The alternative tea market system which started as a small isolated activity has since grown with a substantial number of farmers reporting participation (table 4.2). Considering that marketing arrangements may have impacts on farm operations and activities, this study analysed the effect of tea farmers' participation in ATMC on technical efficiency (TE) in the study area. Understanding the impact of ATMC participation is important for policy on the required institutional arrangements for efficiency improvement.

In order to analyse the effect of ATMC on TE, the study first tested the hypothesis of significant differences in TE between the ATMC participating and non-participating tea farming households. This was achieved using the student t statistic test based on the null hypothesis that there were no statistically significant differences in TE between the ATMC participating and non-participating households. The results of the test are reported in table 4.10.

Table 4.10: Differences in Mean TE Scores and Yields Based on ATMC participation

	Participant (n=191)		**non participants (n=334)**		**Difference**	**statistic**
Variable	Mean	Std dev	Mean	Std dev		t value
TE score	0.50	0.17	0.43	0.13	0.07***	3.26
Tea yields / acre	3154.74	154.73	2512.15	108.98	642.59***	3.46

- Legend: The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- Source: Author's computation

From the result in table 4.10, it can be observed that ATMC participating households had higher average TE scores and mean yields than the non-participating households. This observation provides a hint on the possible effects of ATMC participation on TE. Since causal effects of ATMC participation on TE

would not be ascribed on the basis of the results given that the decision on ATMC participation is endogenous, the study utilized the endogenous switching regression (ESR) model specified in chapter three to estimate the effects. This is because the ESR model can account for self-selection and endogeneity in situations where one wishes to establish the effect of being in one of any two regimes on desired outcomes (Lee, 1982; Kim, *et. al.,* 2000). In this study, the outcome of interest is TE and the two regimes are whether or not a household participated in ATMC or not.

The ESR model consists of a binary equation that links the participation decision with the possible covariates of participation and two separate TE equations; one for the ATMC participants and another for the non-participants. The estimates of the binary model are used to construct the selectivity correction terms (inverse mills ratio) conditional on whether a tea farming household had participated in an ATMC or not. The selectivity correction terms act as instrumental variables in the estimation of bias reduced TE equations (Lokshin and Sajaia, 2004). The estimation of model was achieved using the full information maximum likelihood (FIML) estimation. The FIML was preferred over the two-step least squares or the two-step maximum likelihood estimation since its parameters are consistent, homoscedastic and asymptotically efficient (Maddala, 1986). Further, the two step approach requires potentially cumbersome adjustments to derive consistent standard errors. The FIML Estimates of the ESR model including the results of the diagnostics tests are presented in table 4.11.

Table 4.11: The FIML Estimates of the ESR Model for ATMC Participation

Dependent variable	ATMC Participation (1/0)		TE score (ATMC=1)		TE score (ATMC=0)	
Variable	Coefficient	P value	Coefficient	P value	Coefficient	P value
Farm size	-0.0382	0.638	-0.569**	0.004	-0.108***	0.000
County dummy	-0.390**	0.029	0.154***	0.000	0.141***	0.000
Variety of tea	0.0734	0.573	-0.008	0.885	0.008	0.762
Age of tea farm	0.0361**	0.044	-0.0091	0.117	-0.015***	0.000
Age of tea farm ^{2}	0.0003	0.273	0.000	0.413	0.0002***	0.001
age of farmers	-0.0215***	0.002	0.003	0.160	0.002*	0.098
Education (primary)	0.1590	0.657	0.066	0.606	0.0677	0.306
Education (sec)	0.3147	0.369	0.055	0.664	0.047	0.474
Education (college)	0.4245	0.278	0.018	0.892	0.052	0.510
Extension (FFS)	-0.3391***	0.009	0.124***	0.001	0.025	0.345
Distance to market	-0.0050	0.827	-0.004	0.406	-0.002	0.676
Labour structure	0.600	0.705	0.052	0.282	0.045	0.165
Constant	-0.910	0.862	0.559***	0.000	0.390***	0.000
Gender of H head	-0.1278	0.289				
Residence of H head	0.1446	0.405				
Access to credit	0.2441**	0.037				
Household assets	-0.001*	0.029				
Output of tea (Kgs)	0.001***	0.000				
Food sufficiency	0.323**	0.031				
Per capita expenditure	0.000	0.105				
Rho 0	0.75*	0.076				
Rho 1	0.95**	0.018				
Wald Test	Chi squared = 38.76 (p= 0.0001).					
LR test of independent equations	Chi squared = 50.3 (P=0.000)					

Legend:

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- The default education status is university
- Source: Author's computation

The results reported in the second column of table 4.11 are the binary probit model estimates and corresponds to equation 3.24 in chapter three. The equation provides results on how various factors influence the tea household's decision on ATMC participation. The factors are discussed in section 4.4.1. The results reported in the third and fourth column of table 4.11 correspond respectively to equation 3.34 and 3.35 in chapter three. The results account for the endogenous switching in the TE equations for the two regimes (participants and non-participants). The equations were therefore used to compute the expectations of TE conditional on ATMC participation. The computed conditional expectations of TE were then used to derive the effects of ATMC participation (ATET) based on equation 3.15 also specified in chapter three (3).

After estimation, the model was subjected to various diagnostic tests to check validity of the model. The results of the tests are reported on the bottom part of table 4.11. The Wald test was used to test whether the estimated coefficients jointly estimated as a group were different between the participants and non-participants. The chi squared value of 38.90 (p =0.0001) implies that the coefficients of the two outcome equations reported in column three and four were statistically different. In order to test whether the selection bias adjustment was significant, the study used the likelihood ratio test for joint independence of the three equations. The results of the test rejected the null hypothesis that all coefficients were equal to zero at the 1% level (chi squared value was 64.9 with p= 0.000) and therefore confirmed that the selection bias adjustment was necessary. The covariance correlation coefficients for errors of the binary equation and those

of the continuous equations reported as rho 0 and rho 1 were both non-zero and significant. This implies that both observed and non-observed factors influenced TE and the decision on ATMC participation. The results also confirmed that selectivity biases existed and uncorrected estimates would have been inconsistent.

Based on the results of the diagnostic tests, the estimates of the ESR model were considered to be suitable for meeting the study's objective of determining the effects of ATMC participation on TE.

4.5.1 Factors Influencing Household Participation in ATMC

The variables included in the model for analysing the determinants of ATMC participation were the farm level characteristics (farm size, county dummy, variety of tea planted, age of farm, volume of tea harvested per year and distance to the nearest market), the household demographic characteristics (age of farmer, education, household labour structure, gender, residence of the household head, the household economic characteristic (household assets and per capita expenditure) and institutional variables (access to FFS extension and credit).

The results from the model show that the county dummy, age of farm, age of farmer, access to extension, credit, household assets and household food sufficiency were the important determinants of ATMC participation in the study area. The coefficients of farm size, variety of tea, education, distance to markets, labour structure, gender and residence of the household head and household per capita expenditure were insignificant. This implies that the variables were not important in explaining the farmers' decisions on ATMC participation.

On the farm level characteristics, the coefficient for the county dummy was negative and significant at five percent. This implies that on average a tea farmer in Bomet County was less likely to participate in an ATMC than a farmer in Nyamira County. The coefficient for age of the farm was significant at five percent. This means that on average, a farmer with an older tea farm had a higher probability of participating in an ATMC than a farmer with younger tea. While an a priori explanation for the positive association is less obvious, the relationship could be related to the farmer experience in tea farming. This is because experience broadens the farmer's social network where more market information can be acquired leading to the establishment of more market linkages (Shilpi and Umali-Deininger, 2007).The coefficient for the quadratic term for age of the farm was insignificant, therefore refuting the possibility of a non-linear relationship between age of the farm and participation in ATMC.

On the quantity of tea output, the estimates show that the variable had a positive coefficient which was significant at one percent. The result suggests that an increase in the quantity of tea output harvested by tea farming household was associated with a higher probability of ATMC participation. This result may be explained by the more flexible systems for green leaf collection and delivery that are offered by the ATMC. Compared to KTDA, the alternative markets have less stringent standard for green leaf quality and their schedules for green leaf collection and delivery allow for longer plucking hours or own transportation of green leaf to the factories (Owuor *et.al,* 2005). Higher quantities of tea output would therefore influence the propensity for ATMC participation since farmers

with more tea output are more likely to have capacity for self-transport of green leaf than the tea farmers with less output. Similarly, the farmers with more tea output are more likely to require the convenience offered by the flexibility of the ATMC system than would, the farmers with lower levels of tea output. This observation seems to affirm the finding by Chepkulei (2013), who found that green leaf quality requirements and flexibility in the collection of green tea had a positive influence on participation in ATMC.

Under the household demographic characteristics, the coefficient for age of the farmer was negative and statistically significant at one percent. This suggests that an increase in the age of the farmer was associated with a decline in the probability of participation in ATMC. This observation is consistent with Chepkulei (2013) who found that smallholder tea farmers supplying KTDA factories were on average older than those selling through the alternative markets. The finding is also consistent with Yoshiko (2011) whose study found a preference among older farmers in Vietnam to supply their green leaf to the state owned enterprise factories because of being more averse to market risks.

The coefficient for FFS extension was negative and significant at one percent. This means that the probability of ATMC participation was negatively linked with participation in FFS extension. This finding was not unexpected since the KTDA factories are the leading providers of the FFS extension service to the farmers (Mose *et. al.*, 2016). It is therefore possible that the KTDA extension agents could be integrating messages intended to discourage farmers from ATMC participation

in their extension programs thereby causing the negative association between FFS extension and ATMC participation.

As shown in table 4.11, coefficient for access to credit was positive and statistically significant at five percent. The result suggests that the farmers who had access to credit were more likely to participate in an ATMC than the farmers without access to credit. This relationship may be attributed to the fact that both borrowing and ATMC participation are strategies that households make to relax liquidity constraints connected with the household immediate consumption needs. It is also possible that households which had borrowed may have opted to participate in ATMC as a strategy for circumventing or evading loan deductions through the KTDA check off system.

Under the economic characteristics, the coefficient for household assets was negative and statistically significant at five percent. This implies that the probability of ATMC participation decreased with increase in level of household assets. This result was expected due to the presence of differences in the structure of payment for green leaf adopted by the different market channels. The KTDA factories have a structure for green leaf payment that is characterized by lower monthly rates and higher bonus at the end of the year (Table A.10 in the Appendix). The interpretation is that the households with less household assets have fewer alternative sources of income and are therefore more likely to participate in an ATMC in order to access liquidity to sustain their livelihoods.

With regard to the household food sufficiency, the results show that the variable's coefficient was positive as expected and was statistically significant at five percent. This means that households that relied more on the market to access food staples were more likely to participate in the alternative markets for green leaf than the households that relied on their own production. The variable was modelled as a proxy for demand for liquidity given the substantial share of food expenditure in the budgets of households in the area covered by the study. The observation that households that relied more on the market to access food staples had a higher propensity for ATMC participation affirms the earlier finding that the need for liquidity to meet immediate consumption needs such as food and school fees would be an important motivation for ATMC participation.

Overall, the results of the binary model revealed that the regional (county) fixed effects, age of the farm, the level of tea output, the age of the farmer, FFS extension, access to credit, the level of household assets and household food sufficiency were important determinants of ATMC participation. The influence of farm size, variety of tea planted, education, labour structure and distance to markets was found to be less important in explaining the household decisions on ATMC participation. The other variables that were found to be less important in explaining the household ATMC participation were gender and residence of the household head and per capita household expenditure.

4.5.2 The Effects of ATMC Participation on TE

In order to derive the effects of ATMC participation on TE, the study used the estimates of the two TE equations of the ESR model (column 3 and 4 in table

4.11) to estimate the conditional expectation of TE conditional on ATMC participation. The derivation followed the simulation in equation 3.15 in chapter three and involved obtaining the difference in the expectations of TE conditional on ATMC participation for the participating households and expectations of TE conditional on non-participation for the same households. This was equivalent to obtaining the difference in TE of the farm households that had participated in the ATMC and the TE of the same households had they not participated. The difference is commonly known as average treatment effects of the treated (ATET) in treatment effects literature (Lee, 1982; Rosenbaum and Rubin, 1983; Kim, *et. al.,* 2000; Lokshin and Sajaia, 2004; Awotide *et. al.,* 2015 a; Awotide *et. al.,* 2015b). The results of the simulation reported as ATET is presented in table 4.12.

Table 4.12: The Effects of ATMC Participation from the ESR Model

$E(TE_{1i}\|D_i = 1; X)$		$E(TE_{2i}\|D_i = 1; X)$		$E(TE_{1i}\|D_i = 1; X) - E(TE_{2i}\|D_i = 1; X)$
Mean TE	Std dev	Mean TE	Std dev	ATET
0.49	0.11	0.45	0.13	0.050

Source: Author's computation

Table 4.12 shows the ATET was 0.05 which implied that participation in ATMC had a positive impact on TE. The result implies that participation in ATMC would increase the TE of the farm household by 0.050. The Positive effect of ATMC on TE would be attributed to the more flexible arrangements for green leaf collection and leaf quality. Farmers who sell their green leaf to KTDA factories face more stringent requirements on leaf quality and have got to deliver their green leaf

through designated tea buying centres (TBCs) (Owuor *et. al.,* 2005). Such arrangements involve longer transaction time required to complete a green leaf sale at the TBC (Mbeche, 2012). Unlike KTDA, The ATMC system takes lesser time since green leaf is mainly collected at the farm gate or roadside which therefore minimises the post-harvest losses associated with delayed tea collection. In addition, the differences in transaction time mean that the ATMC participating farm households were more likely to allocate more time to tea farming activities and hence their higher TE than the non-participating farm households.

The positive effect of ATMC participation would also be attributed to the payment structure adopted by the ATMCs which is characterised by higher initial payment rates compared to the rate offered by KTDA. The structure of paying farmers by KTDA provides for higher final (bonus) payments and lower initial monthly payments as shown in Table A.10 in the Appendix. The timing of payment is important because it can affect the frequency and timeliness of farm husbandry practices. At farm level, tea cultivation requires continued and regular financing for inputs and labour (Chirwa and Kydd, 2006). Delayed payment would therefore have a negative effect on the smallholders' ability to fund tea farm operations such as plucking whose expenses constitute a significant proportion of the tea production costs.

In addition, the differences in TE performance may also be attributed to the service packages embodied in ATMC operations. The farmers participating in ATMC have better access to wider industry networks and therefore better information

flow, improved access to production technologies, practices and innovations (Fungo *et. al.,* 2011).

The finding of this study regarding the impact of ATMC participation is consistent with Chirwa and Kydd (2006) who found that the type of market channel through which farmers sold their green leaf was important in explaining variations in tea productivity in Malawi. Specifically, it was found that the smallholder farmers who had contractual arrangements with the commercial estates had higher productivity than those contracted by the reformed state enterprise factory. The finding is however inconsistent with Yoshiko (2011) who found that smallholder tea farmers who were contracted by state owned enterprise (the equivalent KTDA before liberalization) had higher TE levels than the farmers who were not contracted.

The results of the ESR model may be sensitive to its assumptions of joint normality of the error terms. This is because of the sensitivity of the selection bias adjustment to departures from normality (Awotide *et. al.,* 2015 b). The propensity score matching (PSM) model proposed by Rosenbaum and Rubin (1983) was therefore adopted to check the robustness and validity of the estimated effects. PSM provides an alternative approach for dealing with endogeneity biases and works by selecting (matching)a group of non- participants who are similar to the participants of a program in all relevant pre-participation characteristics, and then attributes the differences in the outcomes between the groups to the program or treatment.

The propensity scores used in the matching process were calculated as the probability of participating in an ATMC using the binary probit model (Rosenbaum & Rubin, 1983; Asres *et. al.,* 2013; Shiferaw *et. al.,* 2014). The model whose output is reported in Table A.11 in the Appendix had the same covariates as those in the binary part of the ESR model. After estimation of the propensity scores, matching of the farm household groups was done using the nearest neighbour matching (NNM) algorithm which is commonly applied in PSM studies. In NNM, each ATMC participant is matched to closest non participant using the estimated propensity scores.

After the matching process, the quality of the matching was tested using two procedures that aimed to assess whether the overlap or common support condition which is necessary for the PSM was satisfied in the data (Asres *et. al.,* 2013). This condition assumes that some randomness is required in order to guarantee that farm households with identical characteristics can be observed in both states (Heckman *et. al.,* 1999). The first procedure involved checking the density distribution of the propensity scores after the matching process. A visual inspection of the density distribution shown in Figure A.2 (in the Appendix) indicates that there was substantial overlap in the distribution of propensity scores for both the ATMC participating and non-participating farm households. This therefore led to the conclusion that the common support condition was fulfilled. In addition observations that were outside the region of common support were eliminated from the estimation of participation effects.

The second procedure involved testing the hypothesis that differences in the means of the covariates for both groups were not significant (after matching). Insignificant differences imply that the matching is successful in balancing the distribution of relevant variables in both groups. The results of the balancing test are reported in Table A.12 in the Appendix. The results show that except for household assets, there were no significant differences in the means of the covariates for both for the ATMC participants and ATMC participants. The results revealed that the differences were considerably reduced after matching for all the variables except for household assets, where the difference was not fully eliminated. This suggests that the specification of the propensity scores was fairly successful in balancing the distribution of covariates between the two matched groups and the PSM process was therefore suitable for the estimation of the participation effects.

The effect of ATMC participation on TE was calculated as the difference in average TE between the two matched groups. The results are reported as the average treatment effects for the treated (ATET) in table 4.13.

Table 4.13: The Effects of ATMC Participation from PSM

Estimator	Outcome	Effect	Coefficient	AI Robust SE	Z value	P value
NNM	TE score	ATET	**0.053***	**0.029**	**1.82**	**0.069**

*** denotes significance at 10%** **Source: Author's computation**

***** AI robust standard errors are used to generate heteroskedastic –robust variance estimators to correct for potential heteroskedasticity (Abadie and Imbens, 2002).

The results in table 4.13 show that the coefficient for ATET was 0.053 which suggests that participation in ATMC had a positive and statistically significant effect on TE. On the basis of both the ESR and PSM models, the study therefore found a consistent result indicating that participation in ATMC increases TE. Overall the results provide evidence that the alternative marketing arrangements are important drivers of efficiency and therefore productivity growth within the smallholder tea sub sector in Kenya.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND POLICY IMPLICATIONS

5.1 Introduction

This chapter presents the summary of the study and the conclusions that were derived from the findings. The policy implications, contribution to knowledge and areas for further research are also presented.

5.2 Summary

The tea industry plays an important role in Kenya's economy. The industry is a top foreign exchange earner and indirectly supports over 5 million Kenyans, making it one of the leading sources of livelihood in the country. Globally, Kenya is among the four leading producers; alongside china, India and Sri Lanka, who collectively account for over 75 percent of the global tea production. Tea production in Kenya is dualistic; involving smallholder and large-scale production systems. Although the performance of the smallholder subsector has been characterized by substantial growth in acreage, production and number of growers, productivity has remained low. This prompted an investigation on technical efficiency which is an important factor of productivity growth. The specific objectives of the study were to estimate the level of technical efficiency of smallholder tea production in selected counties in Kenya; determine the effect of farm-size on the technical efficiency of smallholder tea production in selected counties in Kenya; and analyse the effect of green leaf tea marketing arrangements on the technical efficiency of smallholder tea production in selected counties in Kenya.

The study adopted the non-experimental cross-sectional research design to address its objectives. The study targeted a population of 528, 817 smallholder tea farming households in Kenya. Using the multi stage sampling approach, data for the study was collected from a sample of 525 tea farming households selected from Bomet and Nyamira counties of Kenya. In order to estimate the level of TE, the study used the non-parametric DEA model to obtain TE scores for each of the sampled farm households. The scores were then summarised using frequency distributions and means to show the level of TE in the study area. To provide a more complete characterization of efficiency, the study also analysed the distribution of slacks and returns to scale for the tea farms in the sample. The effect of farm size on TE was explored using a fractional regression model which was estimated using the quasi-maximum likelihood estimation technique. The model was considered to be the most ideal for the since it takes into account the fractional nature of efficiency scores In the model, the TE scores obtained from DEA were regressed against the linear and quadratic terms of farm size and other covariates hypothesised to influence TE. The effect of marketing arrangements on TE was analysed using the endogenous switching regression (ESR) model and robustness of the results was checked using the PSM model.

Results of the study revealed that smallholder tea farmers in the study area were technically inefficient. The mean average TE score was 0.46 implying that smallholder tea farmers can achieve their current tea production levels using only 46 percent of their existing resources, if they employ the production practices of the best practice farmers. The level of TE score was 0.55 for Bomet and 0.41 for

Nyamira. Comparatively, the results show that smallholder tea farmers in Nyamira County were technically less efficient than their counterparts in Bomet County. The results further revealed that majority of the smallholder tea farms in the study area were not operating at the optimal scale and technology was characterized by increasing returns to scale.

With regard to the effects of farm size on TE, the study found that farm size had nonlinear effects on TE, with TE first falling and then rising with increase in farm size. The threshold of farm size beyond which increase in farm size leads to a rise in TE was estimated at 3.93 acres. This threshold provides an empirical basis for reforms on farm size in the industry and represents the scale of operation where the advantages of small farms are outweighed by the economies of scale of the larger farms. The study further found that at lower levels of farm size, labour use intensity was important in explaining the relationship between farm size and TE. A part from farm size, the factors that were found to increase TE were access to FFS extension and the share of family labour applied in tea farming. Factors that were found to have a negative effect on TE were the age of the tea farm and the education level of the household head.

In analysing the effect of green leaf tea marketing arrangements on TE, the study considered whether a household had participated in an alternative tea market channel (ATMC) or not. The results show that 36.4 percent of the smallholders utilised the alternative market channels to sell their produce. The factors that were found to increase the likelihood of participation in the ATMC were the level of tea output, access to credit, household demand for liquidity (food sufficiency) and the

age of the tea farm. The factors that negatively influenced the likelihood of ATMC participation were the age of the farmer, access to FFS extension and the level of household assets. In addition, it was found that the county specific dummy accounted for differences in the probability of ATMC participation. With regard to the effect of ATMC participation on TE, the study found that participation in the ATMC increases TE by 0.05 in smallholder tea production in the study area.

5.3 Conclusions

Based on the findings of the study, the following conclusions emerge. First, the study concludes that smallholder tea farmers in the study area were technically inefficient and hence there exists a considerable potential to improve tea productivity in the study area. It is also concluded that by employing the production practices of the best practice farmers, smallholder tea farmers in the study area can achieve their current tea production levels using only 46 percent of their existing resources. Drawing from the existence of differences in TE between tea farm households in Bomet and Nyamira County, the study concludes that regional county specific heterogeneities are important for production efficiency and that such heterogeneities should be accounted for in the formulation of the industry's policies.

Second, the study concludes that farm size has a nonlinear effect on TE and that TE can only increase with farm size for farms beyond the threshold of 3.93 acres. Therefore increasing farm holdings under tea will only be beneficial if farm consolidation programs produce holdings with 3.93 acres or more. The study also concludes that there are labour market imperfections facing smallholder tea

production and that TE increases with the share of family labour used in tea production. Additionally TE increases with increase in the access to FFS extension and reduces with increase in the age of the tea farm and the education level of the household head.

Third, participation in ATMC increases TE, indicating that marketing arrangements are important in tea production. Household participation in ATMC increases with the level of tea output, access to credit, household demand for liquidity (food sufficiency) and the age of the tea garden, but reduces with the age of the farmer, access to FFS extension from KTDA agents and the level of household assets.

5.4 Policy Implications

Drawing from the major findings of the study, some policy implications are suggested. First, the Agricultural and food Authority (AFA) and the county governments should enact and effect regulations that discourage and restrict subdivision of farms below 3.93 acres. The recommendation is supported by the finding that TE was found to increase with farm size for farm sizes above the threshold of 3.93 acres. Since the acreage of most of the tea farms already falls below 3.93 acres, AFA and the respective county governments who are responsible for policy formulation and regulation in the industry should design and promote programs aimed at consolidating the management of small tea farms. A possible model would involve strategies that encourage the smallholder farmers to release their farms to individual or corporate professional farm managers for collective management and joint performance of production activities. Such a

model would allow the collectivised farms to benefit from the scale advantages and increased efficiency.

The Tea Research Institute (TRI), KTDA and other industry players should develop and promote labour saving technologies so as to address labour sourcing bottlenecks within the subsector. The recommendation derives from the finding that at lower levels of farm size (< 3.93 acres) TE was found to decline with farm size due to differences in labour use intensity. In addition, focus should be given towards correcting the imperfections in the labour markets in the tea sector. This is because the increase in the share of family labour applied in tea farming was found to increase TE; which implies that hired labour is an imperfect substitute for family labour and hence the presence of imperfections in the labour markets.

AFA and the county governments should ensure that the policy formulation in the tea sector accounts for the county specific heterogeneities in the different tea growing counties. This would require measures that foster adequate participation of all the key stakeholders at the county level during the policy formulation process. The recommendation derives from the fact that agriculture in a devolved function and is also supported by the finding that county specific heterogeneities accounted for differences in TE in tea production in the study area.

AFA and the county governments should implement tea replanting programs to replace the aging or moribund tea gardens. This is in light of the finding that increase in age of the tea farm has a negative effect on TE. The mean age of tea farms in the study area was 27.0 years which implies that tea gardens in the study

area are relatively aged and therefore need to be replaced. Since tea replanting involves colossal amounts of money due to the high costs of tea establishment and forgone incomes, the county governments should liaise with the national government and other donors for funding. Possible implementation options would involve the provision of replanting subsidies or low interest loans to the farmers.

KTDA and the other players in the tea sector should expand the FFS extension program to replace the conventional Train and Visit (T&V) extension system which is faced with many challenges such as poor linkages between research and extension and higher ratio of farmers -to- extension agent (Musa *et. al.*, 2013). The findings of the study show that FFS extension had a positive effect on TE and that only 52.8 percent of the sample had participated in the FFS program. Enlisting more farmers in the program is important since extension services is one of the critical change agents in the tea sector given its role in the dissemination of knowledge, information and technologies and linking farmers with other actors in the value chain.

Finally AFA and the county governments should expand market reforms in order to increase the farmers' access to the alternative green leaf markets. This is in light of the finding that participation in ATMC has a positive effect on TE. Liberalization of the tea industry should be reviewed in order to fully liberalize the sub-sector and allow for more competition in the market. The review should also aim at addressing outstanding challenges so as to foster an enabling environment that will allow entry and participation of more competitors in the green leaf tea market segment.

5.5 Contribution to Knowledge

This study adds to the existing literature on efficiency and provides greater insights on the sources of inefficiency in the smallholder tea subsector in Kenya. It is the first study that analysed the effects of farm size on TE in the Kenyan sector. A part from contributing to the debate on the relationship between farm size and efficiency that exists within the broader agricultural economics literature, the study explicitly demonstrates that there exists a threshold of farm size beyond which increase in farm size leads to an increase TE. The study therefore provides an empirical basis for policy reforms on farm size in the Kenyan tea sector. Previous studies in the sector had analysed efficiency in the context of differences in agro-ecological conditions. This study analysed efficiency in the context of two counties which are relatively similar in agro-ecological conditions and therefore demonstrate that differences in TE can also be due to fixed county heterogeneities.

The introduction of liberalization reforms in smallholder tea sub sector was intended to spur efficiency through increased accountability and competition. However there has been a lack of empirical evidence linking the ATMCs which are an outcome of the reforms with efficiency. This is the first study that empirically analysed the effect of ATMC participation on TE. The study demonstrates that participation in ATMC increases TE, therefore pointing to the need to deepen market reforms in the smallholder tea subsector. The study also contributes to the understanding of the various factors that motivate participation in the ATMCs which is a critical input for policy.

5.6 Areas of Further Research

The following areas are identified for further research;

i) The effect of ATMC participation on profitability in tea production among smallholder tea farmers, since participation in ATMC has costs and benefits.

ii) An empirical analysis on the determinants of household labour allocation and sources of labour market imperfections in the smallholder tea subsector.

iii) An empirical investigation on how green leaf pricing and payment structure adopted by the different players affects productivity in the smallholder tea subsector.

iv) While the study provides insights on the need for farm size reforms that are applicable within the wider industry context, there is therefore need for more empirical studies on the subject to cover the other tea growing counties not covered by the study.

REFERENCES

Abadie, A. & Imbens, G. (2002). Simple and Bias-Corrected Matching Estimators for Average Treatment Effects. Technical Working Paper T0283, National Bureau of Economic Research, Inc.

Aigner, D. J., Lovell, C. A. K. & Schmidt, P. (1977). Formulation and Estimation Of Stochastic Frontier Production Function Models, *Journal of Econometrics*, Vol. 6:21-37.

Alene, A., Manyong, V., Omanya, G., Mignouna, H., Bokanga, M. & Odhiambo, G. (2008). Smallholder Market Participation Under Transactions Costs: Maize Supply and Fertilizer Demand in Kenya, *Food Policy*, Vol. 33:318–328.

Antonakis, J., Bendahan, S., Jacquart, P. & Lalive, R. (2010). On Making Causal Claims: A Review and Recommendations, *The Leadership Quarterly*, 21(6): 1086–1120.

Asefa, S. (2011). Analysis of Technical Efficiency of Crop Producing Smallholder Farmers in Tigray, Ethiopia. Munich Personal RePEc Archive (MPRA) Paper No. 40461, Mekelle University.

Awotide, B.A., Alene, A.D. & Abdoulaye, T. (2015a). Impact of Agricultural Technology Adoption on Asset Ownership: The Case of Improved Cassava Varieties in Nigeria, *Food Security*, Vol. 7 (6):1239–1258.

Awotide B.A., T. Abdoulaye , Alene, A. & Manyong, V. M. (2015b). Impact of Access to Credit on Agricultural Productivity: Evidence from Smallholder Cassava Farmers in Nigeria. International Conference of Agricultural Economists: Milan, Italy August 9-14, 2015.

Asres E., N. Makoto., Kumi, Y. & Akira, I. (2013). Effect of Agricultural Extension Program on Smallholders' Farm Productivity: Evidence from Three Peasant Associations in the Highlands of Ethiopia, *Journal of Agricultural Science*; Vol. 5 (8): 164-181.

Bhalla, S. & Roy, P. (1988). Misspecification in Farm Productivity Analysis: The Role of Land Quality, *Oxford Economic Papers*, 40(1): 55-78.

Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies, *Management Science*, 39, 1261-1264.

Banker, R. D., Cooper, W. W., Seiford, M. L., Thrall, R. M. & Zhu, J. (2004). Returns to Scale in Different DEA Models, *European Journal of Operational Research*, Vol. 154:345–362.

Barrett, C. B. (1996). On Price Risk and the Inverse Farm Size-Productivity Relationship, *Journal of Development Economics*, Vol. 51 (2): 193-215.

Battese, G. E & Coelli, T. (1988). Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data, *Journal of Econometrics*, Vol. 38 (3): 387-399.

Basnayake B.M.J.K. & Gunaratne, L.H.P. (2002). Estimation of Technical Efficiency and It's Determinants in the Tea Small Holding Sector in the Mid Country Wet Zone of Sri Lanka, *Sri Lankan Journal of Agricultural Economics*, 4 (1): 137-150.

Baten M. A., Kamil, A. A. & Haque, M. A. (2010). Productive Efficiency of Tea Industry: A Stochastic Frontier Approach, *African Journal of Biotechnology*, Vol. 9(25): 3808-3816.

Becker, G. S. (1965). A Theory of the Allocation of Time, *Economic Journal*, Vol. 75 (299): 493-517.

Bore, J.K. (1996). A Review of Problems of Old Tea Fields, *Tea Journal*, Vol. (17), 27-33.

Caliendo, M. & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching, *Journal of Economic Surveys*, Vol. 22 (1): 31-172.

Cameron, C. & Trivedi, P. (2005). *Micro-econometrics: Methods and Applications*, Cambridge: Cambridge University Press.

Carrasco, R. (2001). Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labor Participation, *Journal of Business & Economic Statistics*, 19(4): 385–94.

Carter, M. R. & Wiebe, K. D. (1990): Access to Capital and its Impact on Agrarian Structure and Productivity in Kenya, *American Journal of Agricultural Economics*, 72(5): 1156-1150.

Chamber, R. G. (1988). *Applied Production Analysis*, Cambridge: Cambridge University Press.

Chand R., Prasanna, L. & Singh, A. (2011). Farm Size and Productivity: Understanding the Strengths of Smallholders and Improving Their Livelihoods, *Review of Agriculture. Economic and Political Weekly Supplement*, Vol. XIVI (26) & 27 (5).

Charnes, A., Cooper, W. W. & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, Vol. 2: 429–444.

Chayanov, A.V. (1966). *The Theory of Peasant Economy,* in ed. Daniel Thoner. Homewood, Ill. Richard, D. Irwin, Hunt.

Chepkulei, B. (2013). Determinants of Market Choice among Tea Farmers. Proceedings of the 8th JKUAT Scientific, Technological and Industrialization Conference, November 15, 2013.

Chirwa, E. W. & Kydd, J. (2006). Farm-Level Productivity in Smallholder Tea Farming in Malawi: Do Contractual Arrangements Matter? Working Paper No. 2006/03. University of Malawi Chancellor College: Department of Economics.

Clougherty, J. & Duso, T. (2015). Correcting for Self-Selection Based Endogeneity in Management Research: A Review and Empirical Demonstration. Discussion Papers of DIW Berlin 1465, DIW Berlin, German Institute for Economic Research.

Coelli, T.J. (1995). Recent Developments in Frontier Modelling and Efficiency Measurement, *Australian Journal of Agricultural Economics,* Vol. 39(3): 219-245.

Coelli T. J. (1996). A Guide to DEAP Version 2.1: A Data Envelopment Analysis Program. Centre for Efficiency and Productivity Analysis Working paper, Department of Econometrics, University of New England.

Coelli, T. J. & Perelman, S. (1996). A Comparison of Parametric and Nonparametric Distance Functions with Application to European Railways. CREPP Discussion paper, Liege: University of Liege.

Coelli, T. J., Rao, D. S. P., Donnell, C. J. O. & Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis*, 2nd ed. New York: Springer.

Cooper, W., Seiford L. M., Tone, K & Zhu, J. (2007). Some Models and Measures for Evaluating Performances with DEA: Past Accomplishments and Future Prospects, *Journal of Productivity Analysis*, Vol. (28): 151-163.

Daraio, C. & Simar, L. (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, New York: Springer.

Debertin, D. L. (2012). *Agricultural Production Economics, 2nd Eds*. New Jersey: Macmillan Publishing Company.

Debreu, G. (1951). The Coefficient of Resource Utilization, *Econometrica*, Vol. 19, (3):273-292

de Janvry, A., Fafchamps, M. & Sadoulet, E. (1991). Peasant Household Behavior with Missing Markets: Some Paradoxes Explained, *Economic Journal*, 101(409): 1400-17.

de Janvry, A. & Sadoulet, E (2007). Progress in the Modeling of Rural Households Behavior under Market Failures. In Alain de Janvry and Ravi Kanbur (ed.) Poverty, Inequality and Development. *Essays in Honor of Erik Thorbecke*: Kluwer publishing.

Desbiez, A., Matthews, R., Tripathi, B. & Ellis J. (2005). Perceptions and Assessment of Soil Fertility by Farmers in the Mid-hills of Nepal, *Agriculture, Ecosystem and Environment,* Vol. 10 (1):191-206.

Dinh, C. N. (2011). Performance of Small and Medium Enterprises and the Impact of Environmental Variables: Evidence from Vietnam. Unpublished PhD Thesis, Birmingham: Aston University, U.K.

Dyer, G. (1991). Farm Size–Farm Productivity Re-Examined: Evidence From Rural Egypt, *Journal of Peasant Studies,* Vol. 19(1): 59–92.

Ellis, F. (1992). *Peasant Economics*: *Farm Household and Agrarian Development*, Cambridge: Cambridge University Press.

Fafchamps, M. (1992). Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World, *American Journal of Agricultural Economics*, Vol. 74(1):90- 99.

Fafchamps, M. & Vargas, R. H. (2005). Selling at the Farm-gate or Travelling to Market, *American Journal of Agricultural Economics,* 87(3): 717-734.

Färe, R., Grosskopf, S. & Lovell, C. A. K. (1994). *Production Frontiers*, Cambridge: Cambridge University Press.

Farrell, M. J. (1957). The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, Series A (General), Vol. 120 (3): 253-290.

Fungo B. , Clark L. , Tenywa M. M. , Tukahirwa J., Kamugisha R. , Birachi E., Wanjiku C., Bizoza A. R. , Wimba B. , Pali P. , Adewale A. & Olowole F.(2011). Networks among Agricultural Stakeholders in the Southwestern Highlands of Uganda, *Journal of Agricultural Extension and Rural Development*, Vol. 3(7):118-129.

Gesimba, R.M., Langat, M.C., Liu G. & Wolukau, J.N. (2005). The Tea Industry in Kenya; The Challenges and positive Developments, *Journal of Applied Sciences,* 5 (2):334-336.

Gelan Ayele & Muriithi, B. W (2015). Examining Returns to Scale in Smallholder Dairy Farms in East Africa. *Quarterly Journal of International Agriculture*, Vol. 3: 239-261.

Gilligan, D. O. (1998). Farm Size, Productivity, and Economic Efficiency: Accounting for Differences in Efficiency of Farms by Size in Honduras. Paper Presented at 1998 Annual meeting of the American Agricultural Economics Association, Salt Lake City, Utah, August 2-5.

Goetz, S. J. (1992). A Selectivity Model of Household Food Marketing Behavior In Sub-Saharan Africa, *American Journal of Agricultural Economics*, Vol. 74(2):444-52.

Gorton, M. & Davidova, S. (2001). Farm Productivity and Efficiency in the CEE Applicant Countries, *Agricultural Economics*, 30: 1–16.

Govereh, J., Jayne T. S. & Nyoro, J. (1999). Smallholder Commercialization, Interlinked Markets and Food Crop Productivity: Cross-Country Evidence in Eastern and Southern Africa. Unpublished, Department of Agricultural Economics and Department of Economics, Michigan State University.

Greene, W. (2003). Simulated Likelihood Estimation of the Normal-Gamma Stochastic Frontier Function, *Journal of Productivity Analysis*, Vol. (19):179–190.

Gunther, S. (1995). Why is the Agriculture of Advanced Western Economies Organized by Family Farms? Will this continue to be so in Future? Theme of Plenary Paper *No. 2: in Agricultural Economics. The International Library of Critical Writing in Economics,* Edited By George H. Papers, Edward Elgar Publishing company: 443-458.

Guo, S. & Fraser, M. W. (2010). *Propensity Score Analysis: Statistical Methods And Applications,* Thousand Oaks, CA: Sage Publications.

Hall, P. (1994). Methodology and Theory for the Bootstrap", in: R. E. Engle and D. McFadden, (eds), *Handbook of Econometrics*, Vol. 4, Amsterdam: Elsevier.

Heady, E. O. & Dillon, J. L. (1969). *Agricultural Production Functions,* Ames, Iowa: Iowa State University press.

Heath, H. (2015). Considering Technical and Allocative Efficiency in the Inverse Farm Size-Productivity Relationship, *Journal of Agricultural Economics*, Vol. 66(2): 442-469.

Heckman, J. J. (1974). Shadow prices, market wages, and labor supply, *Econometrica*, 42(4), 679-694.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error, *Econometrica*, 47(1), 153-161.

Heckman, J., LaLonde, R. & Smith, J. (1999): The Economics and Econometrics of Active Labor Market Programs," *in Handbook of Labor Economics* Vol.III, ed. by O. Ashenfelter, and D. Card, Amsterdam: Elsevier,

Helfand, S. M. & Levine, E. S. (2004). Farm Size and the Determinants of Productive Efficiency in the Brazilian Center-West, *Agricultural Economics*, Vol. 31:241-249.

Hong, N. B., Takahashi Y. & Yabe, M. (2015). Resource Use Efficiency of Tea Production in Vietnam: Using Translog SFA Model, *Journal of Agricultural Science*, Vol. 7 (9).

Huffman, W. E. (2010). Working Paper No. 10019 June 2010. Household Production Theory and Models. Working Paper No. 10019, June 2010, Ames, Iowa: Department of Economics, Iowa State University.

International Tea Committee (2013). Annual Bulletin of Statistics, Colombo: Aitken Spence Printing (Pvt) Ltd, Sri Lanka.

Jayne, T. S. & Muyanga, M. (2012). Land Constraints in Kenya's Densely Populated Rural Areas: Implications for Food Policy and Institutional Reform. Department of Agricultural, Food, and Resource Economics, Michigan State University. Contributed Paper prepared for presentation at the 86th Annual Conference of the Agricultural Economics Society, University of Warwick, United Kingdom - 18 April 2012.

Jehle, G. A. & Reny, P. J (2011). *Advanced Microeconomic Theory*, 3rd Ed, Edinburgh Gate: Pearson Education Ltd.

Ji, Y. & Lee, C.(2010). Data Envelopment Analysis, *The Stata Journal 10*, Vol. (2):267–280.

Jondrow, J., Lovell, C., Materov, I. & Schmidt, P. (1982). On the Estimation of Technical Efficiency in the Stochastic Production Function Model, *Journal of Econometrics*, Vol. 19:233-238.

Just, R. & Pope, R. (2001). *Handbook of Agricultural Economics, Vol. 1*. Ed B Gardner and G Rausser , North Holland: Elsevier science.

Kagira E. K., Kimani, S. W. & Githii, K. S. (2012). Sustainable Methods of Addressing Challenges Facing Small Holder Tea Sector in Kenya: A Supply Chain Management Approach, *Journal of Management and Sustainability*; Vol. 2 (2): 75-89.

Kamau, D. M. (2008). Understanding Smallholder Tea Farmers; Closing the Loop Between Expectations and Realities, *Tea Journal*, Vol. 29 (2): 25-29.

Karagiannis, G. (2014). Modeling Issues in Applied Efficiency Analysis: Agriculture, *Economics and Business Letters*, Vol. 3(1): 12-18.

Kavoi, M. M., Owuor, P .O., Siele, D. K. & Oluoch-Kosura, W. (2001). A Test For Relative Efficiency in the Smallholder Tea Sub-sector in Kenya, *Journal of Agriculture, Science and Technology*, Vol. 3(1): 22-29.

Kiani, A. K. (2008). Farm Size and Productivity in Pakistan, *European Journal of Social Sciences, Vol.* 7 (2): 42-52.

Kim, S., Nayga, R. M. & Capps, O. J. (2000). The Effect of Food Label Use on Nutrient Intakes: An Endogenous Switching Regression Analysis, *Journal of Agricultural and Resource Economics;* Vol. 25(1):215-231.

KIPPRA (2013). Kenya Economic Report 2013 on Creating an Enabling Environment for Stimulating Investment for Competitive and Sustainable Counties. Nairobi: Kenya Institute for Public Policy Research and Analysis.

Kiprono, P. K. (2013). Economic Efficiency of Resource Use among Smallholder Tea Producers in Selected Counties of Kenya, Unpublished PhD Thesis: Moi University.

Kinyili, J. M. (2003). Diagnostic Study of the Tea Industry in Kenya, Nairobi, Export Promotion Council.

Koopmans, T. (1951). *Activity Analysis of Production and Allocation*, New york: John Wiley & Sons.

Kothari, R.C. (2004). *Research methodology, Methods and Techniques*, New Delhi: New International (P) Ltd, publishers.

KTDA (1963-2017). *KTDA Annual Reports and Accounts*, Various Issues (1963-2017), Nairobi:, KTDA, Kenya.

KTDA (2010). *Strategic Business Plan 2009 – 2014*, Nairobi: KTDA, Kenya.

Koskei, R.C. Langat, J.K., Koskei, E.C. & Oyugi, M .A. (2013). Determinants Of Agricultural Information Access by Small Holder Tea Farmers in Bureti District, Kenya, *Asian Journal of Agricultural Sciences,* 5(5): 102-107.

Kumar, S. & Gulati, R. (2008). An Examination of Technical, Pure Technical, And Scale Efficiencies in Indian Public Sector Banks using Data Envelopment Analysis, *Eurasian Journal of Business and Economics*, Vol. 1 (2), 33-69.

Lee, L. F. (1978). Unionism and Wage Rates: a Simulation Equation Model with Qualitative and Limited Dependent Variable, *International Economic Review,* Vol. 19: 415-455.

Leonard, D. K. (1991). *African Successes: Four Public Managers of Kenyan Rural Development*, Berkeley: University of California Press.

Lokshin, M. & Sajaia, Z. (2004). Maximum Likelihood Estimation of Endogenous Switching Regression Models, *Stata Journal*, Vol. 4(3), 282-289.

Lovell, C. A. K. (1993). *Production Frontiers and Productive Efficiency, in the Measurement of Productive Efficiency*: *Techniques and Applications* (Eds.) H. Fried, C.A.K. Lovell and S. Schmidt, Oxford : Oxford University Press.

Maddala, G. S. (1986). Disequilibrium, Self-Selection, and Switching Models. In: Griliches and Intriligator, *Handbook of Econometrics*, Vol. 3:1633-1688.

Maity, S. (2012). Farm size and economic efficiency: A Case Study of Tea Production in West Bengal, *International Journal of Sustainable Economy*, Vol. 4(1): 53–70.

Maro G, Mrema J., Msanya, B. & Teri, J. (2013). Farmers' Perception of Soil Fertility Problems and their Attitudes towards Integrated Soil Fertility Management for Coffee in Northern Tanzania, *Journal of Soil Science and Environmental Management*, Vol. (5):93-99.

Marschak, J. (1960). Binary-Choice Constraints and Random Utility Indicators (1960). Arrow Kenneth, eds *Stanford Symposium in Mathematical Methods in Social Sciences*, Stanford: Stanford University Press.

Masterson, T. (2007). Productivity, Technical Efficiency and Farm Size in

Paraguayan Agriculture. Working Paper No. 490: The Levy Economics Institute of Bard College.

Mbeche, R. (2012). The Impact of Liberalization Policies and other initiatives on Downward Accountability and Empowerment of small scale Tea Farmers in Kenya, Unpublished PhD Thesis, University of Reading, UK.

McDonald, J. (2009). Using Least squares and Tobit in Second Stage DEA Efficiency Analyses, *European Journal of Operational Research*, Vol. 197(2): 792-798.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (eds.), *Frontiers of Econometrics*, (pp 105-142), New York: Academic Press.

McFadden, D. (2000). Economics of Choices. Nobel Lecture. Department of Economics University of California, Berkeley, CA 94720-3880 USA.

Mendola, M. (2007). Farm Household Production Theories: A Review of "Institutional" and "Behavioral" Responses, *Asian Development Review*, Vol. 24(1): 49-68.

Meeusen, W. & van den Broeck, J. (1977). Efficiency Estimation from Cobb–Douglas Production Function with Composed Error. International, *Economic Review*, Vol. 18: 435–444.

Mose G. N., Mbeche R. & Ateka, J. (2016). Institutional Innovations for Smallholder Agricultural Production Systems in Kenya: A Case of Smallholder Tea Subsector, *European Journal of Sustainable Development* (2016), **5**, 3, 461-475.

Mugera, W., Amin, K. & Langemeier, M. R. (2011). Does Farm Size and Specialization Matter for Productive Efficiency? Results from Kansas, *Journal of Agricultural and Applied Economics, Vol* 43, (4):515–528.

Muniu J., Mburu T. & Obere , A. (2015). *The value Effects of Bank Mergers and Acquisitions in Kenya,* LAP Lambert Academic Publishing.

Musa, Y.N, Aboki, E. & Audu, I. (2013. The Limitations and Implications of Training and Visit (T&V) Extension System in Nigeria, *Journal of Agriculture and Sustainability,* Vol. 4(1): 67-76

Mwaura, F.M., Nyabundi, K. & Muku, O. (2005). Situation Analysis of the Small–Scale Tea Growers and Their Contribution at the Local Auction Market in Kenya, *Tea Journal,* Vol. (2):35-45.

Mwakha, E. (1989). Rehabilitation of Moribund Tea Progress Report, *Tea Journal,* Vol. 10: 124-133.

Mwaura, F. M., Muku, O., Marangu, D. & Towett E. (2008). Technological and Socio–Eeconomic Factors Affecting Tea Productivity among Smallholders in Imenti and Kapkoros, *Tea Journal* Vol. 29 (2):19-24.

Nguyen-Van, P. & To-The, N. (2014). Agricultural Extension and Efficiency of Tea Production in Northeastern Vietnam, *Bureau deconomie theorique et appliquee (Beta),* 7522.

Njeri, E. K., Kimani, E. N. & Lodiaga, J. M. (2013). Gender Dynamics in the Access and Control of Benefits Accrued from Tea Farming in Kiganjo Division, Gatundu District, *Applied Science Reports*, Vol. (1): 6-10.

Nyangena, W. (2008). Social Determinants of Soil and Water Conservation in Rural Kenya, *Environment, Development and Sustainability*, 10:745–767, DOI 10.1007/s10668-007-9083-6.

Nyangito, H. & Kimura, J. (1999). Liberalization of the Smallholder Tea Sub–sector; Progress Impacts and Recommendations for Further Development. Discussion paper No. 016 of 1999, Nairobi: Institute of policy Analysis and Research.

Nyangito, H. O. (2000). Policy and Legal Framework of the Tea Subsector and the Impact of Liberalization in Kenya. Policy Paper No. 1, Nairobi: Kenya Institute of Public Policy Research and Analysis.

Obare G. A., Nyagaka, D. O. Nguyo, W. & Mwakubo, S. M. (2010). Are Kenyan Smallholders Allocatively Efficient? Evidence from Irish Potato Producers in Nyandarua North District, *Journal of Development and Agricultural Economics*, Vol. 2(3): 078-085.

Ogada M. J., Muchai, D., Mwabu G. & Mathenge, M. (2014).Technical Efficiency of Kenya's Smallholder Food Crop Farmers: Do Environmental Factors Matter? Journal of *Environment, Development and Sustainability; a Multidisciplinary Approach to the Theory and Practice of Sustainable Development*; Vol. 16 (5).

Ogise, M., Mwaura, M. F., Nyabundi, W. K. & Ateka, J. M. (2008). Incorporating Farmers Characteristics and Aspirations in Poverty Reduction Strategies among Tea Smallholders in Kenya. *Tea Journal,* Vol. 29 (1): 23-30.

Ogundari, K., Ojo, S. O. & Ajibefun, I. A. (2006). Economies of Scale and Cost Efficiency in Small Scale Maize Production: Empirical Evidence from Nigeria. *Journal of Social Science,* Vol. 13(2): 131-136.

Omiti, J. M. & Mosoti, A. (2010). Position Papers on Appropriate Tea Industry Legal and Institutional Framework, Multiple Taxation and Value Addition, October 2010, Nairobi: East African Tea Association (EATTA).

Omiti, J., Otieno, D., Nyanamba, T. & Mc Culloug, E. (2009). Factors Influencing the Intensity of Market Participation by Smallholder Farmers: A Case Study of Rural and Peri-urban Areas of Kenya, *African Journal of Agricultural and Resource Economics*, Vol. 3 (1):1-26.

Onduru, D.D., De Jager, A., Hiller, S. &Van den Bosch, R. (2012). Sustainability of Smallholder Tea Production in Developing Countries: Learning Experiences from Farmer Field Schools in Kenya, *International Journal of Development and Sustainability*, Vol. 1 (3): 714-742.

Owuor, P.O. (1999). Tea in East Africa, in *Global Advances in Effect of Maceration Method on Chemical Tea Science*, New Delhi: New Age International Ltd, India.

Owuor, P.O., Kavoi, M.M. & Siele, D.K. (2005). Assessment of Constraints in Technology Transfer and Policies which Limit the Realization of High Green Leaf Production in the Smallholder Tea Sector of the Kenya Tea Industry: An Empirical Analysis of Economic Efficiency and Supply of Tea. Research Paper No. 19, Nairobi: Africa Technology Policy Studies.

Padilla-Fernandez, M.D.& Nuthall, P. L. (2012). Farm Size and Its Effect on the Productive Efficiency of Sugar Cane Farms in Central Negros, the Philippines, *Journal of the International society of Southeast Asian Association of Agricultural sciences*, Vol. 18(1): 49-61.

Papke, L.E. & Wooldridge, J.M. (1996). Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. *Journal of Applied Econometrics*, Vol. 11(6): 619-63.

Pascoe, S., Parastoo, H., Jesper, A. & Korsbrekke, K. (2001). Economic Versus Physical Input Measures in the Analysis of Technical Efficiency in Fisheries. Paper presented at the XII Conference of the European Association of Fisheries Economists, Salerno, Italy, 18-20 April 2001.

Pareto, V. (1909). *Manual of Political Economy*, Paris: Giars & Briere.

Quisumbing, A. R. (1996). Male-Female Differences in Agricultural Productivity: Methodological Issues and Empirical Evidence, *World Development, Vol.* 24 (10): 1579-1595.

Ramalho, E.A., Ramalho, J.J.S. & Henriques, P.D. (2010). Fractional Regression Models for Second Stage DEA Efficiency Analyses, *Journal of Productivity Analysis*, Vol. 34: 239-255.

Ramalho, E.A., Ramalho, J.J.S. & Murteira, J. M. R. (2011). Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models, *Journal of Economic Surveys*, Vol. 25(1): 19-68.

Republic of Kenya (1999). *Sessional Paper No. 2 on the Liberalization and Restructuring of the Tea Industry: Reforms in the Tea Board of Kenya and Privatization of the Kenya Tea Development Authority*, Nairobi: Government Printer.

Republic of Kenya (2005-06). *Kenya Integrated Household Budget Survey (KIHBS) 2005-06*. Nairobi: Government Printers.

Republic of Kenya (2007). *The Task Force Report on the Tea Industry in Kenya*. Nairobi: Government Printer.

Republic of Kenya, (2008). *Kenya Vision 2030*, Nairobi: Government Printer.

Republic of Kenya (2009). *The 2009 Kenya Population and Housing Census 2009,* Nairobi: Government Printers.

Republic of Kenya (2010). *Agricultural Sector Development Strategy 2010–2020,* Nairobi: Government Printers.

Republic of Kenya (2011). *The Tea (Amendment) Act, 2010*, Nairobi: Government Printers.

Republic of Kenya (2013). *The Agriculture, Fisheries and Food Authority Act of of 2013*, Nairobi: Government Printers .

Republic of Kenya (2014a). *The Draft National Tea Policy,* Nairobi: Government Printers.

Republic of Kenya (2014b). *Kenya Demographic and Health Survey of 2014,* Nairobi: Government Printers.

Republic of Kenya (2015). *Economic Review of Agriculture [ERA] 2015*, Nairobi: Government Printers.

Republic of Kenya (2016). *The Task Force Report on the Tea Industry in Kenya*. Nairobi: Agriculture and Food Authority.

Republic of Kenya (2017). *Kenya Economic Survey 2017*, Nairobi: Government Printers.

Rios, A. R. & Shively, G. E. (2005). Farm size and Nonparametric Efficiency Measurements for Coffee Farms in Vietnam, Department of Agricultural Economics, Purdue University, presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27.

Rosenbaum, D. & Rubin, B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects Paul R. *Biometrika*, Vol. 70 (1): 41-55.

Rosenzweig, M. R. (1980). Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Market Family Labor Supply in a Developing Country Author(s): *The Quarterly Journal of Economics*, Vol. 94 (1). 31-55.

Sanne, van D. (2008). Sustainability Issues in the Tea Sector: A Comparative Analysis of Six Leading Producing Countries. Amsterdam: SOMO Centre for Research on Multinational Corporations, the Netherlands.

Schultz, T. W. (1964). *Transforming Traditional Agriculture*, Chicago: University of Chicago Press.

Seiford LM & Zhu, J. (1999). An investigation of Returns to Scale in Data Envelopment Analysis. *Omega, International Journal of Management Science*, Vol. 2: 1-11.

Shiferaw B., Menale, K., Moti, J. & Chilot, Y. (2014). Adoption of Improved Wheat Varieties and Impacts on Household Food Security in Ethiopia, *Food Policy*, Vol. 44:272–284.

Shilpi, F. & Umali-Deininger, D. (2007). Where to Sell? Market Facilities and Agricultural Marketing, Washington DC: Policy Research Working Paper, No. 4455, World Bank.

Simar, L. & Wilson, P. (2007). Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes, *Journal of Econometrics*, Vol. 136: 31-64.

Singh, I., Squire L. & Strauss, J. (1986). *Agricultural Household Models: Extensions, Applications and Policy*, Baltimore: Johns Hopkins University Press.

Taylor, E. & Adelman, I. (2003). Agricultural Household Models: Genesis, Evolution and Extensions, *Review of Economics of the Household*, 1:33-84.

Tea Board of Kenya (2012). *Annual Report 2011*, Nairobi: Tea Board of Kenya.

Tea Board of Kenya (2013). *Annual Report 2012*, Nairobi: Tea Board of Kenya.

Tea Board of Kenya (2014). *Annual Report 2013*, Nairobi: Tea Board of Kenya.

Thurston, A. (1987). *Smallholder Agriculture in Colonial Kenya: The Official Mind and the Swynnerton Plan*, Cambridge: Cambridge African Monographs 8, African Studies Centre.

Tolga T., N. Yildiz , Nargeleçekenler, M. & Çetin, B. (2009). Measuring the Region, Turkey, *New Zealand Journal of Crop and Horticultural Science*, Vol: 37 (2): 121-129.

TRFK (2002). Tea Growers Handbook, 5th ed, Kericho: Tea Research Institute, Kenya.

Varian, H.R. (1992). *Microeconomic Analysis*, 3rd ed., New York: W.W. Norton Company.

Vincova, K. (2005). Using DEA Models to Measure Efficiency, *BIATEC*, Vol. XIII(8).

Verbeek, M.(2012). *A Guide to modern Econometrics*, 4th ed., New York: John Wiley & Sons Inc.

Githinji, M., Konstantinidis C. & Barenberg, A. (2011) . Small and as Productive: Female Headed Households and the Inverse Relationship between Land Size and Output in Kenya, UMASS Amherst Economics Working Papers 2011-31, University of Massachusetts Amherst: Department of Economics.

Wickramaratne, M. R. T. (1981). Genotype-Environment Interactions in Tea (Camellia Sinensis) and their Implication in Tea Breeding and Selection, *Journal of Agricultural Science (Cambridge),* Vol. 96:471-478.

Wooldridge, J. (2002). *Econometric Analysis of Cross-Section and Panel Data*. Massachusetts: MIT Press.

Wooldridge, J. (2010). *Econometric Analysis of Cross-Section and Panel Data*. 2nd eds., Massachusetts: MIT Press.

Yoshiko, S. (2011). Contract Farming and its Impact on Production Efficiency and Rural Household Income in the Vietnamese Tea Sector. Unpublished PhD Thesis, University of Hohenheim: Institute of Agricultural Economics and Social Sciences in the Tropics and Subtropics.

Zhu, J. (2014). *Quantitative Models for Performance Evaluation and Benchmarking, International Series in Operations Research & Management Science,* 3rd ed., Switzerland: Springer International Publishing.

APPENDICES

7.1 Appendix 1: Tables And Figures

Table A.1: Tea Farming Households and Acreage across Counties in Kenya

County	No of farm households	Production kgs (3-year average)	Planted area (ha)	Productivity Kgs PT/Ha	Ha/farm
Kiambu	24,604	36,576,142.00	11,046.00	2224	0.45
Muranga	62,237	37,835,799.70	14,286.50	2236	0.23
Nyeri	31,652	18,718,980.30	7,501.50	2396	0.24
Kirinyaga	39,705	15,930,472.00	7,235.00	2643	0.18
Embu	23,515	10,454,299.70	4,306.80	2332	0.18
Meru	55,117	26,862,525.30	11,445.20	2236	0.21
Kericho	50,322	78,566,109.30	38,836.10	1566	0.77
Bomet	46,154	45,555,728.00	26,100.60	1748	0.57
Nyamira	87,034	31,022,457.30	17,690.40	1620	0.2
Kisii	74,623	16,310,993.70	12,090.90	1435	0.16
Nandi	13,080	54,278,274.00	4,930.00	1570	0.38
Trans Nzoia	1,970	1,325,321.30	485.9	2306	0.25
Kakamega	15,158	2,896,771.00	2,161.00	1348	0.14
Nakuru	3,646	5,301,236.30	2,041.80	2435	0.56
Total/avarage	528,817	381,635,110.00	160,157.90	1748	0.3

The figures are for 2010-2013
Source: TBK (2013)

Table A.2: Potential Productivity of Tea Clones Released by TRI

Clone	Source of seed	Yields (Kg Pt/ ha)	Year released	Quality index[$]
TRFK 303/577	OP of TRFK 6/8	7817	1989	Medium quality
TRFK 371/8	OP of AHP S15/10	6641	2012/13	High quality
TRFK 371/3	OP of AHP S15/10	6000	2008	High quality
TRFK 303/178	OP of TRFK 6/8	5722	1986	Medium quality
TRFK 303/1199	OP of TRFK 6/8	5569	1989	Medium quality
TRFK 303/216	OP of TRFK 6/8	5383	1986	Plain quality
TRFK 303/231	OP of TRFK 6/8	5286	1989	Plain quality
TRFK 31/8	Ambangulu, Tanzania	5049	1964	Medium quality
TRFK 6/8	Kericho, Kenya	4441	1964	high quality
TRFK 303/156	OP of TRFK 6/8	4410	1994	high quality
TRFK 303/259	OP of TRFK 6/8	4351	1988	Plain quality
TRFK 337/3	OP of TRFK 6/8	4104	1995	Medium quality
TRFK 31/27	Ambangulu, Tanzania	4100	1988	Medium quality
TRFK 337/138	OP of TRFK 6/8	4097	1995	Plain quality
TRFK 338/13	OP of TRFK 31/11	4097	1995	poor quality
TRFK 306	OP of TRFK 91/1	4000	2011	High quality
TRFK 303/999	OP of TRFK 6/8	3945	1989	Medium quality
TRFK 303/791	OP of TRFK 6/8	3927	1989	Plain quality
TRFK 303/978	OP of TRFK 6/8	3920	1989	Plain quality
TRFK 303/199	OP of TRFK 6/8	3840	1988	Plain quality
TRFK 303/186	OP of TRFK 6/8	3798	1994	poor quality
TRFK 31/11	Ambangulu, Tanzania	3753	1969	high quality
TRFK 303/152	OP of TRFK 6/8	3591	1994	Medium quality
TRFK 303/179	OP of TRFK 6/8	3591	1994	Medium quality
TRFK 303/745	OP of TRFK 6/8	3523	1989	poor quality

Source TRI: (1989-2013)

**Total polyphenols (catechins) were estimated from fresh shoots by spectrophotometric method, following which the relative quality of the clones can be described as: high quality (24.80 - 27.07%), medium high quality (22.47 - 24.40%), medium quality (19.57 - 22.33%) and low quality (17.53 - 19.17%).* [$] *Quality indices were derived from Tea Tasters' sensory evaluation data as determined at the Mombasa tea auction.*

Table A.3: Estimated TE Scores from DEA

ID	TE	ID	TE	ID	TE	ID	TE	ID	TE	ID	TE
1	0.44	51	0.28	101	0.36	151	0.40	201	0.25	251	0.60
2	0.36	52	0.20	102	0.90	152	0.19	202	0.15	252	0.20
3	0.62	53	0.18	103	0.22	153	0.43	203	0.11	253	0.26
4	0.50	54	0.07	104	0.19	154	0.26	204	0.91	254	0.61
5	0.57	55	0.47	105	0.40	155	0.41	205	0.50	255	0.49
6	0.26	56	1.00	106	0.10	156	0.66	206	0.43	256	0.25
7	0.21	57	0.38	107	0.18	157	0.28	207	0.52	257	0.27
8	0.41	58	0.28	108	0.42	158	0.66	208	0.75	258	0.35
9	0.20	59	0.30	109	0.33	159	1.00	209	0.44	259	0.67
10	0.29	60	0.11	110	0.26	160	0.66	210	0.60	260	0.39
11	0.23	61	0.42	111	0.66	161	0.24	211	0.39	261	0.26
12	0.11	62	0.24	112	0.42	162	1.00	212	0.24	262	0.42
13	0.19	63	0.43	113	0.59	163	0.56	213	0.60	263	0.29
14	0.21	64	0.07	114	0.18	164	0.22	214	0.66	264	0.48
15	0.53	65	0.08	115	0.27	165	0.74	215	0.29	265	0.48
16	0.30	66	0.36	116	0.44	166	1.00	216	0.56	266	0.16
17	0.19	67	0.31	117	0.23	167	0.49	217	0.59	267	1.00
18	0.15	68	0.61	118	0.42	168	0.23	218	0.38	268	0.80
19	0.36	69	0.29	119	0.48	169	0.26	219	0.41	269	0.16
20	0.20	70	0.35	120	0.27	170	0.25	220	0.21	270	0.31
21	0.30	71	0.40	121	0.11	171	0.72	221	0.66	271	0.35
22	0.11	72	0.45	122	0.36	172	0.76	222	0.46	272	0.22
23	0.91	73	0.24	123	0.37	173	0.24	223	0.66	273	0.46
24	0.41	74	0.55	124	0.13	174	0.21	224	0.24	274	0.39
25	0.31	75	0.37	125	0.40	175	0.14	225	0.33	275	0.20
26	0.36	76	0.17	126	0.14	176	0.12	226	1.00	276	0.13
27	0.23	77	0.56	127	0.16	177	0.33	227	0.27	277	0.13
28	0.37	78	0.12	128	0.50	178	0.43	228	0.36	278	0.42
29	0.34	79	0.77	129	1.00	179	0.46	229	1.00	279	0.44
30	0.14	80	0.42	130	0.26	180	0.53	230	0.21	280	0.84
31	0.26	81	0.36	131	0.25	181	0.19	231	0.43	281	0.30
32	0.28	82	0.46	132	0.22	182	0.23	232	1.00	282	0.47
33	0.33	83	0.71	133	0.24	183	0.41	233	1.00	283	0.27
34	0.33	84	0.39	134	0.31	184	0.16	234	0.12	284	0.71
35	0.34	85	0.19	135	0.18	185	0.33	235	0.18	285	1.00
36	0.37	86	0.35	136	0.37	186	0.40	236	0.21	286	1.00
37	0.15	87	0.51	137	1.00	187	0.16	237	0.33	287	0.34
38	0.15	88	0.33	138	0.32	188	0.28	238	0.80	288	0.62
39	0.28	89	0.82	139	0.19	189	0.29	239	0.26	289	0.22
40	0.17	90	0.60	140	0.74	190	0.43	240	0.40	290	0.45
41	0.40	91	0.28	141	0.36	191	0.39	241	0.24	291	0.26
42	0.21	92	0.17	142	0.31	192	0.77	242	0.36	292	0.42
43	0.29	93	0.44	143	0.18	193	0.63	243	0.43	293	0.24
44	0.26	94	0.16	144	0.47	194	0.25	244	0.36	294	0.72
45	0.19	95	0.11	145	0.54	195	0.23	245	0.72	295	0.49
46	0.69	96	0.34	146	0.44	196	0.35	246	0.43	296	0.72
47	0.33	97	0.36	147	0.54	197	0.18	247	0.23	297	0.78
48	0.49	98	0.39	148	0.40	198	0.32	248	0.50	298	0.98
49	0.32	99	0.68	149	0.27	199	0.25	249	0.11	299	0.41
50	0.26	100	0.38	150	0.23	200	0.15	250	0.19	300	0.44

Table A.3: Estimated TE Scores Continued

ID	TE	HID	TE	ID	TE	ID	TE	ID	TE
301	0.31	351	0.64	401	0.11	451	0.62	501	0.64
302	0.83	352	0.35	402	0.48	452	0.74	502	0.48
303	0.43	353	0.52	403	0.75	453	0.49	503	0.62
304	0.57	354	0.46	404	0.60	454	0.15	504	0.79
305	0.35	355	0.40	405	0.53	455	0.71	505	0.43
306	0.25	356	0.48	406	0.27	456	0.40	506	0.25
307	0.42	357	0.76	407	0.54	457	0.42	507	0.47
308	0.33	358	0.31	408	0.42	458	0.80	508	0.82
309	0.68	359	0.32	409	0.39	459	0.38	509	0.31
310	0.46	360	0.19	410	0.47	460	0.53	510	0.31
311	0.58	361	0.77	411	0.42	461	0.46	511	0.50
312	0.15	362	0.29	412	0.30	462	0.46	512	0.51
313	0.42	363	0.22	413	0.41	463	0.32	513	0.44
314	0.49	364	0.47	414	0.64	464	0.60	514	0.59
315	0.49	365	0.20	415	0.43	465	0.39	515	0.68
316	0.69	366	0.25	416	0.65	466	0.33	516	1.00
317	0.48	367	0.35	417	1.00	467	0.74	517	0.83
318	0.80	368	0.16	418	0.43	468	0.65	518	1.00
319	1.00	369	0.41	419	0.54	469	0.49	519	0.65
320	0.43	370	0.90	420	0.55	470	0.85	520	0.97
321	0.60	371	0.43	421	0.61	471	0.55	521	0.93
322	1.00	372	0.33	422	1.00	472	1.00	522	0.74
323	0.58	373	1.00	423	0.87	473	0.91	523	0.68
324	0.37	374	0.40	424	0.77	474	0.40	524	1.00
325	0.51	375	0.34	425	0.54	475	0.43	525	1.00
326	0.26	376	0.41	426	0.17	476	0.37		
327	0.29	377	0.60	427	0.67	477	1.00		
328	1.00	378	0.54	428	1.00	478	0.62		
329	0.22	379	0.48	429	1.00	479	0.52		
330	0.46	380	1.00	430	0.55	480	0.27		
331	0.41	381	0.53	431	0.41	481	0.75		
332	0.30	382	0.32	432	0.45	482	0.84		
333	0.25	383	0.63	433	0.44	483	0.59		
334	1.00	384	1.00	434	0.38	484	0.29		
335	0.42	385	0.27	435	0.65	485	0.29		
336	0.18	386	0.67	436	0.70	486	0.87		
337	0.70	387	0.25	437	0.57	487	0.34		
338	0.69	388	0.70	438	0.21	488	0.31		
339	0.69	389	1.00	439	0.41	489	0.43		
340	1.00	390	0.72	440	0.48	490	0.29		
341	0.90	391	1.00	441	0.42	491	0.33		
342	0.70	392	1.00	442	0.18	492	0.61		
343	1.00	393	0.39	443	0.50	493	0.52		
344	0.23	394	0.46	444	0.26	494	0.35		
345	0.56	395	0.54	445	0.17	495	0.51		
346	0.59	396	0.53	446	0.40	496	1.00		
347	0.25	397	0.44	447	0.26	497	0.53		
348	0.36	398	0.67	448	0.19	498	0.44		
349	0.62	399	1.00	449	0.77	499	0.52		
350	0.59	400	0.48	450	0.41	500	1.00		

Source: Author's computation
Label : Household IDs
1-331 are for Nyamira County Household IDs
332-525 are for Bomet County

Table A.4: Correlation Matrix for the FR Model Variables

	County Dummy	Household type	Variety	Market channel	FFS extension	Education primary	Education secondary	Education college	Farm size	Labour structure	Distance to market	Age of farm
County Dummy	1											
Household type	0.0505	1										
Variety	0.1337*	0.0512	1									
Market channel	-0.1183*	-0.0347	-0.0194	1								
FFS extension	0.0367	0.071	0.0218	0.0537	1							
Education primary	0.0242	-0.1822*	0.0406	0.0365	-0.1423*	1						
Education secondary	-0.0641	0.089*	-0.0048	-0.0423	0.0831	-0.7465*	1					
Education college	0.0743	0.0872*	-0.053	-0.0514	0.047	-0.2955*	-0.266*	1				
Farm size	0.1654*	0.0351	-0.0444	-0.0883*	-0.0043	-0.0021	-0.0753	0.0542	1			
Labour structure	0.0701	-0.0033	0.1438*	-0.0497	0.1367*	0.092*	0.091*	-0.1428*	-0.3098*	1		
Distance to market	-0.017	-0.0983*	0.1224*	0.0015	0.0147	0.1581*	-0.0994*	-0.085	0.0154	0.0447	1	
Age of farm	-0.082	-0.0745	-0.1082*	-0.006	-0.092*	0.3202*	-0.208*	-0.1384*	0.2099*	-0.096*	0.0447	1

*** denotes significance at 5%** **Source: Author's computation**

Table A.5: Estimation output and Marginal Effects of the FR Model

Wald chi^2(14 df) 200.49 (P=0.000) Log Pseudo-likelihood= -344.67				
Dependent variable: TE score **Number of observations** (n= 525)				
Variable	**Coefficient**	**Std. error**	**Z value**	**P value**
Farm size	-0.416***	0.079	-5.27	0.000
Square of farm size	0.054***	0.016	3.35	0.001
County dummy	0.398***	0.055	7.3	0.000
Variety of tea	0.028	0.049	0.58	0.562
Age of farm	-0.016**	0.007	-2.23	0.026
square of age of farm	0.000*	0.000	1.76	0.078
Household type	-0.038	0.070	-0.55	0.584
Education (primary)	0.203**	0.088	2.29	0.022
Education (secondary)	0.186**	0.086	2.17	0.030
Education (college)	0.193*	0.101	1.9	0.057
Extension (FFS)	0.134***	0.048	2.79	0.005
Distance to market	-0.006	0.011	-0.53	0.598
Labour structure	0.116*	0.061	1.92	0.055
Market channel	0.185***	0.052	-3.58	0.000
Constant	0.155	0.140	1.1	0.271
Marginal Effects				
Farm size	-0.165***	0.031	-5.27	0.000
Square of farm size	0.021***	0.006	3.35	0.001
County dummy	0.157***	0.021	7.37	0.000
Variety of tea	0.011	0.019	0.58	0.561
Age of farm	-0.006**	0.003	-2.23	0.026
square of age of farm	0.000*	0.000	1.76	0.078
Gender of household head	-0.015	0.028	-0.55	0.585
Education (primary)	0.080**	0.035	2.29	0.022
Education (secondary)	0.074**	0.034	2.17	0.030
Education (college)	0.077*	0.040	1.91	0.057
Extension (FFS)	0.053***	0.019	2.79	0.005
Distance to market	-0.002	0.004	-0.53	0.598
Labour structure	0.046*	0.024	1.91	0.055
Market channel	0.07362***	0.021	-3.58	0.000

Legend:

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- Marginal effects for dummy variables refer to the discrete change from 0 to **1**
- The default education status is university
- **Source**: Author's computation

Table A.6: VIF Test for Multi-collinearity for the FR Model

Variable	VIF	Tolerance	R-Squared
Farm size	7.77	0.13	0.87
Square of farm size	7.22	0.14	0.86
County dummy	1.10	0.91	0.09
Variety of tea	1.07	0.93	0.07
Age of tea farmer	1.24	0.81	0.19
Household type	1.06	0.94	0.06
Education (Primary)	6.36	0.16	0.84
Education (secondary)	5.91	0.17	0.83
Education (college)	2.72	0.37	0.63
Extension (FFS)	1.07	0.94	0.06
Distance to market	1.06	0.94	0.06
Labour structure	1.28	0.78	0.22
Market channel	1.04	0.96	0.04
Mean VIF	2.99		

Legend: Critical limit (multicollinearity is a problem if VIF >= 10)

- Source: Author's computation

Table A.7: Marginal Effects of the FR Model with Bootstrapped Std Errors

Variable	Coefficient	Std. Error	t value	P value
Farm size	-0.165***	0.029	-5.64	0.000
Square of farm size	0.021***	0.006	3.54	0.000
County dummy	0.157***	0.025	6.27	0.000
Variety of tea	0.011	0.018	0.62	0.536
Age of farm	-0.006**	0.002	-2.56	0.011
Square of age of farm	0.000**	0.000	2.06	0.040
Household type	-0.015	0.030	-0.5	0.617
Education (primary)	0.080**	0.033	2.47	0.013
Education (secondary)	0.074**	0.034	2.16	0.031
Education (college)	0.077*	0.041	1.88	0.060
Extension (FFS)	0.053**	0.022	2.45	0.014
Distance to market	-0.002	0.003	-0.67	0.504
Labour structure	0.046**	0.023	1.98	0.048
Market channel	0.074***	0.019	-3.92	0.000
Wald : Chi2 276.1 (14) (P=0.000) No of observations: (n= 525)				
Log Pseudo likelihood= -344.6 Replications = 50				

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- Marginal effects for dummy variables refer to the discrete change from 0 to **1**
- The default education status is university
- Source: Author's computation

Table A.8: Output of the Recomputed FR Model (DEA scores without farm size)

Dependent Variable : TE				
Variable	Coefficient	Std. Error	t value	P value
Farm size	-0.153***	0.031	-5.01	0.000
Square of farm size	0.020***	0.006	3.11	0.002
County dummy	0.156***	0.021	7.44	0.000
Variety of tea	0.018	0.019	0.92	0.357
Age of farm	-0.006**	0.003	-2.15	0.031
Square of age of farm	0.000	0.000	1.64	0.102
Household type	-0.012	0.028	-0.43	0.665
Education (primary)	0.069*	0.036	1.9	0.057
Education (secondary)	0.055	0.035	1.55	0.121
Education (college)	0.068*	0.041	1.65	0.099
Extension (FFS)	0.065***	0.019	3.43	0.001
Distance to market	-0.002	0.004	-0.58	0.559
Labour structure	0.032	0.023	1.36	0.175
Market channel	0.066***	0.020	-3.25	0.001
Wald : Chi^2 180.3 (14)	(P=0.000)	No of observations: (n= 525)		
Log pseudo likelihood	= -346.14454			

Legend:

- Source: Author's computation
- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- The default education status is university

Table A.9: Comparison of Farm Size for plots with GSF and those without GSF

Household category	Frequency	Percent	Mean farm size (acres)	t value	P value
Plots with good soil fertility	259	49.3	1.47	2.69	0.0073
Plots without good soil fertility	266	50.7	1.21		

Source: Author's computation

Table A.10: Green Leaf Tea Prices by selected Tea Industry Players in Kenya (2013)

Factory	Initial (Kshs/Kg)	Final payment (Kshs/Kg)	Total (Kshs/Kg)
KTDA	14	31.65	46
Eastern Produce.	20	24	44
Maramba	17	27	44
Kapchorua W N	20	23	43
Koisagat D L	20	19	39
Mabrokie	20	19	39
Tindiret	20	18	38
Unilever	23	10	33
Kaisugu	20	5	25
Sotik Tea – Sasini	21	4	25
Kapchebet	21	3	24

Source: KTDA, 2013

Table A.11: Estimation Output of the Probit Model for PSM

Dependent Variable: **ATMC participation**				
Variable	**Coefficient**	**Std. error**	**Z value**	**P value**
County dummy	-0.241	0.169	-1.42	0.154
Age of tea farmer	-0.022***	0.008	-2.93	0.003
Household type	-0.094	0.189	-0.50	0.618
Age of farm	0.017**	0.008	2.20	0.028
Residence of Hh	0.448	0.250	1.80	0.073
Access to credit	0.671***	0.161	4.17	0.000
Distance to market	-0.007	0.024	-0.30	0.765
Labour structure	0.177	0.169	1.05	0.295
Extension (FFS)	-0.449***	0.137	-3.28	0.001
Output	0.000***	0.000	4.45	0.000
Education (Primary)	0.245	0.415	0.59	0.555
Education (secondary)	0.432	0.406	1.06	0.288
Education (college)	0.808*	0.447	1.81	0.071
Education (university)	0.000	(omitted)		
Food sufficiency	-0.039*	0.023	-1.66	0.096
Per-capita expenditure	0.000	0.000	1.08	0.281
Household assets	0.000**	0.000	-2.43	0.015
Constant	-0.414	0.638	-0.65	0.517

LR chi ^{2} (18) 86.20 (0.000) **source: Author's computation**

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)
- The default education status is university

Table A.12: Balancing Test of Matched Groups (PSM)

Variable	Mean		P value
	Participants	participants	
County dummy	0.43	0.50	0.14
Age of tea farmer	47.45	46.82	0.66
Household type	0.85	0.87	0.65
Age of farm	22.18	22.22	0.98
Household size	6.28	6.41	0.65
Residence of household head	0.96	0.96	1.00
Access to credit	0.82	0.84	0.58
Distance to market	2.90	2.53	0.23
Labor structure	0.63	0.61	0.68
Extension (FFS)	0.49	0.61	0.20
Output	3,945.80	4391.30	0.24
Education (Primary)	0.43	0.43	1.00
Education (secondary)	0.43	0.45	0.75
Education (college)	0.11	0.11	1.00
Education (university)	0.02	0.01	0.18
Farm size	1.41	1.51	0.40
Per-capita expenditure	42,401.00	42719.00	0.93
Household assets	100,000.00	120,000.00	0.087*

- The asterisks denotes significance (*** at 1%, ** at 5% and * at 10 %)

Source: Author's computation

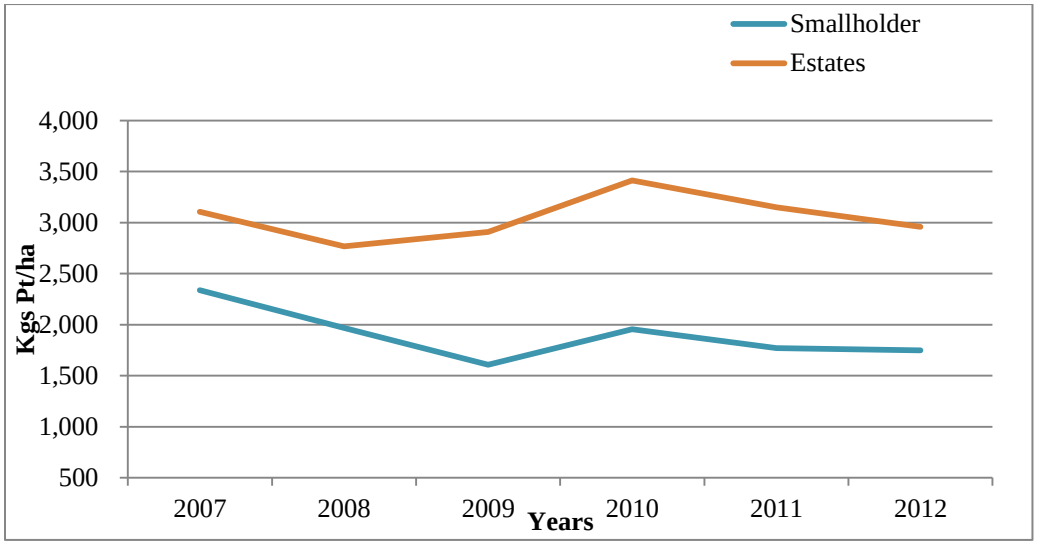


Figure A.1: Smallholder and plantation Tea Productivity (2003-2012)

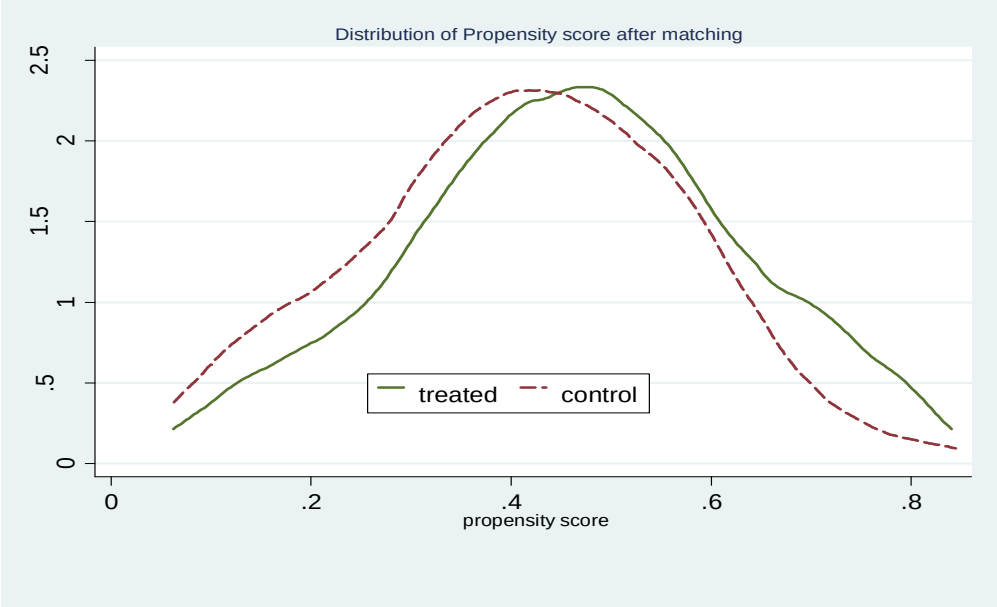


Figure A.2: Distribution of Propensity Scores after Matching

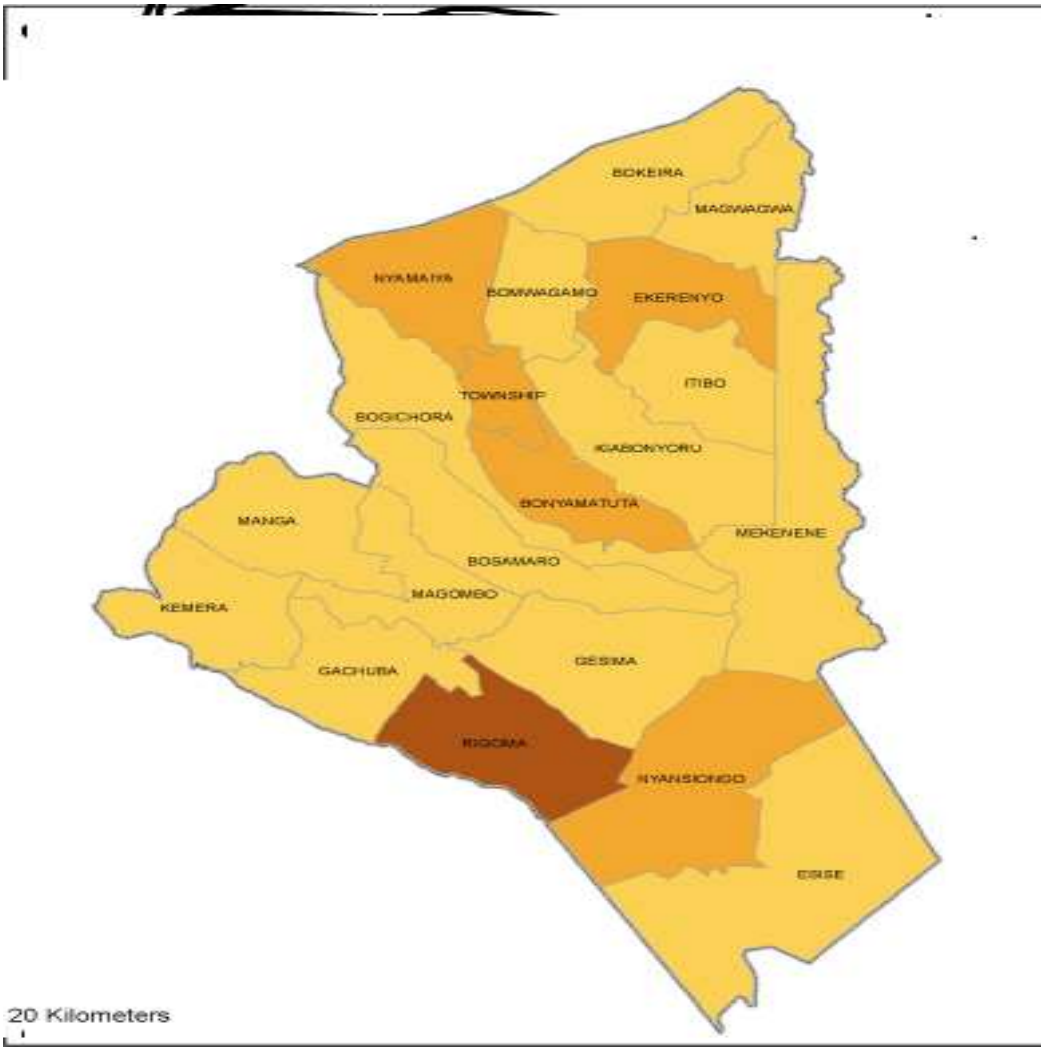


Figure A.3: Nyamira County Map

Source: Nyamira County Government

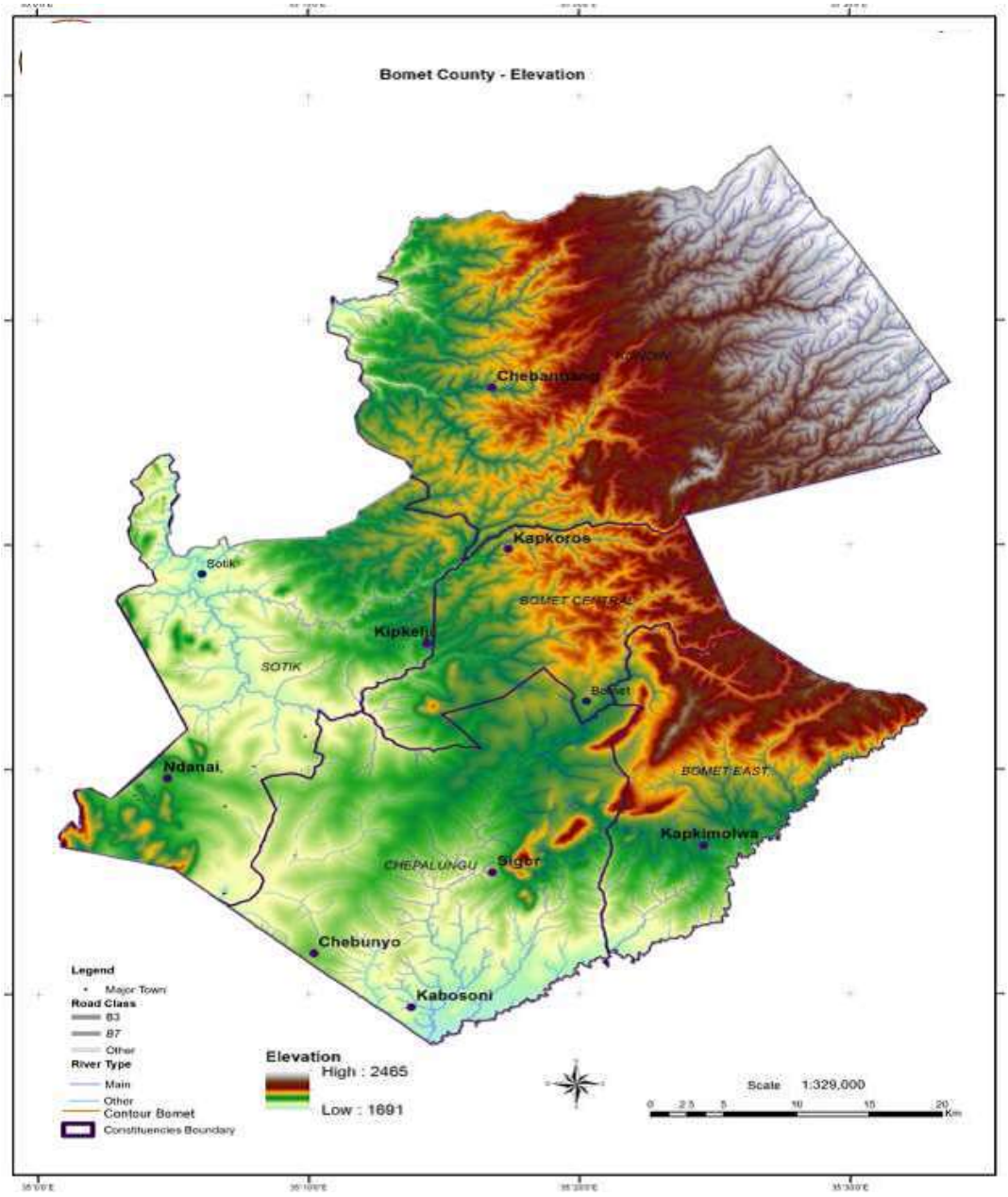


Figure A.4: Bomet County Map

Source: World Wide Fund for Nature (WWF)

7.2 Appendix 2: Profiles of The Selected Study Counties

This study was conducted in Bomet and Nyamira counties of Kenya. The two counties are located in western Kenya and are among the leading tea producing counties in Kenya as shown in Table A.1 in Appendix A.1. The counties share a common boundary and are members of the Lake Region Economic Block (LREB) that brings together 13 counties drawn from part of South Rift and the former Nyanza and Western provinces. The profiles of the counties are described in section 3.8.1 and 3.8.2.

7.2.1 Bomet County

Bomet county borders Kericho county to the North and North East, Narok county to the South East, South, and South West, Nyamira county to the North West and Nakuru county to the East. The County lies between latitudes 0º 29' and 1º 03′ South and between longitudes 35º 05′ and 35º 35′ East and covers an estimated area of 2037.7 Km^2 of land, 85% of which is suitable for crop and livestock production (County Government of Bomet, 2018). Administratively, the county is divided into five sub-counties; namely, Bomet Central, Bomet East, Chepalungu, Sotik and Konoin (County Government of Bomet, 2013).

The demographic indicators in the county are characterized by a high population density and growth rate (2.3%), based on the Kenya National Housing and Population Census Report of 2009 (Republic of Kenya, 2009). The county's population was estimated to be 782,531 in 2012 and was projected to rise to 891,168 by 2017. The Population density was 384 per KM^2 in 2012 and was expected to grow to 437 persons per KM^2 in 2017. Due to the rising population,

farm sizes in the county are relatively small, averaging about 1.5 Ha per household (County Government of Bomet, 2013).

Most parts of the county are characterized by good agricultural potential which is favoured by fertile volcanic soils and well-distributed rainfall that averages 1200–1800 mm per annum. Majority of the farmers in the county practice mixed farming with tea and dairy being the dominant enterprises. Despite the good agricultural potential, the county is characterized by high levels of food insecurity and poverty. The county's has a poverty index of 55 percent against the national average of 47 percent (Republic of Kenya, 2014).

7.2.2 Nyamira County

Nyamira County is located in the Western part of Kenya in the former Nyanza Province and borders Kisii County to the West, Bomet County to the South East, Kericho County to the East and Homabay County to the North. The county lies between latitude 00 30' and 00 45' South and between longitude 340 45' and 350 00' East and covers approximately 899 Km^2. Administratively, the county is divided into five (5) sub-counties namely; Nyamira South, Nyamira North, Borabu, Manga and Masaba North. The county is further sub-divided into 14 divisions with 38 locations and 90 sub locations (County government of Nyamira, 2013).

Demographically, the county is among the most densely populated areas in Kenya with households in most parts of the county living closely adjacent to each other. According to Kenya's National population census of 2009; the county had a population of 598,252 persons and a population density of 665 persons per Km^2

(Republic of Kenya, 2009). The population was projected to rise to 692,641 persons based on the population growth rate of 1.83 percent (County government of Nyamira, 2013).

Nyamira County has a bimodal annual rainfall distribution with the rainfall amounts ranging between 1200–2100 mm per annum. The long and short rain seasons start from December to June and July to November respectively, with no distinct dry spell separating them. Due to the reliability of precipitation and generally good soils, most parts of the county are suitable for agricultural production with the dominant enterprises being tea, maize, banana, sugarcane, dairy and other livestock enterprises. The county is characterized by relatively hilly topography which favours the growth of tea which is the main cash crop and source livelihood in the county. Despite the good agricultural potential, poverty level in the county is high estimated at about 48.8 percent (Republic of Kenya, 2011).

7.3 Appendix 3: Research Questionnaire

Introduction

I am a researcher from Kenyatta University, School of Economics; undertaking a study that seeks to understand the effects of farm size and marketing arrangements on efficiency in the smallholder tea sub sector in Kenya. As part of the research, I am conducting a survey of smallholder tea farmers in Kenya, and you have been identified as one of the respondents. Your participation is entirely voluntary and I do hope that you will agree to participate. The survey will involve you in a semi-structured interview/discussion with some members of our research team. The findings of the research will be written up in research reports which will be appropriately disseminated to stakeholders. I wish to assure you that your responses will be treated with utmost confidentiality and will only be used for the intended purpose.

.

1) General Identification

Item	Response	Item	Response
Questionnaire number		Name County	
Name of grower		Name Sub county	
Date of interview		Name Division	
Name of the enumerator		Name Buying centre	
Name of factory registered in		Registration No	
Telephone		Name Village	

2) Household demographic and social economic characteristics

Question	Indicator	Response
1. Sex of household head	1=Male; 2=Female	
2. Age in years or year born; household head	Actual number of years	
3. Age in years or year born; spouse	Actual number of years	
4. Marital status of household head	1=Married; 2=single; 3=Divorced; 4=Widowed 5=others (specify)	
5. Level of education attained by household head	0 = pre-school 1=Primary ; 2= secondary 3= tertiary 4= university	
6. Level of education attained by spouse	0= pre-school 1=Primary ; 2= secondary 3= tertiary 4= university	
7. How many members are in the household	Actual number	
8. How many members are currently living with you? (Those currently considered members of the household including relatives, workers)	Adult (F+M) aged 60+	
	Adult females (18-59)	
	Adult males (18-59)	
	Children (7-17)	
	Young children below 6 years	
9. Of the adult members in 8 above, How	Number of members	

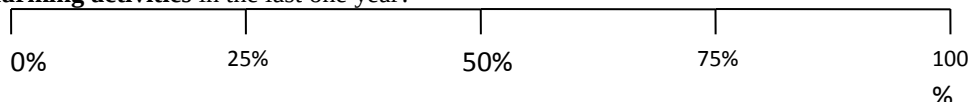
many have at least secondary education		
10. Does the household head reside within the homestead on full time basis?	1=yes; 0= No	
11. If no, where else does he/she reside?	1=within homestead; 2= town or other village; 99=not applicable	
12. On average how many days in a month is the household head available to make decisions on farming	Actual Number of days	
13. What is the main occupation of the household head?	1=farming; 2=teacher; 3=Agric. officer; 4=others (specify)	
14. What is occupation of the spouse?	1=farming; 2=teacher; 3=Agric. officer; 4=others (specify); 99=not applicable	
15. Is tea the main source of income for the household	1=YES; 0= No	
16. How many years have you been farming tea as an independent farmer?	Number of years	
17. Are you a member of any cooperative society or farmer group	1= yes 0= No	
18. If yes to 14 above specify the type	1= Dairy or agricultural society 2= SACCO 3= informal self-help group 4 = other specify	
19. Do you have access to any financial facilities (credit or savings)	1= Yes; 0 = No	
20. Did the household obtain any credit / loan that was used to finance tea farming activities in last financial year (2014/15)	1=YES; 0= No	
21. If no to 17 above why?	1= household did not seek/apply for loan 2=Application was declined (specify reason)	
22. What is the distance from the homestead to nearest market	Actual distance in kilometres	
23. What is the distance to nearest paved road	Actual distance in kilometres	
24. What is the distance to the nearest Tea Buying centre	Actual kilometres	
25. How do you rate the household access to markets for agricultural produce (non-tea)	1= good ; 2 = fair ; 3 = poor	

3) Time Allocation and Household Decision Making

Enumerator Instructions: Introduce this section as follows: "As a farmer you may or may not be involved in other non-farming activities such as business or salaried employment."

a) Is the household involved in any other activities apart from farming
Yes [1] No [2]

b) If yes to 3(a) above what Proportion of **time** did the **Household Head** allocate to the **non-farming activities** in the last one year?



c) Who **makes most of the decisions** on farming for each of the following farm enterprises

No	Enterprise	Code/ indicator	Decision maker
1	**Tea**	(1= self 2= spouse 3=both head and spouse 4=other, specify)	
2	**Maize**	(1= self 2= spouse 3=both head and spouse 4=other, specify	
3	**Dairy**	(1= self 2= spouse 3=both head and spouse 4=other, specify	
4	**Others (specify)**	(1= self 2= spouse 3=both head and spouse 4=other, specify	

4) Land use (use table below)

a) Please use table below to indicate the acreage of parcels operated/ farmed in the last crop season (year)

Item	**No of parcels**	**Total Acres**
How many parcels of land does the household own?		
How many parcels of land did the household lease in last 1 year?		
How many parcels of land did the household rent out in last 1 year (owned but not farmed by you)?		
Total land area farmed or operated in the last year		

b) If the land owned was inherited; what was the initial household heads landholding before sub-division?

Acreage	

c) Kindly provide listing of the main key farm enterprises undertaken and how the land operated is allocated to these enterprises (2014/2015)

	Enterprise	Acres	%		enterprise	acres	%
1	Tea			5			
2	Maize			6			
3	Dairy			7			
4							
	Total						

Section B: Tea Farming

5) Tea production (2014/15)

a) How would you describe the fertility of the main soil on which your tea is planted? (based on physical properties such soil depth, drainage, texture, colour and presence of weeds)

[1] Good [0] Poor

b) How do you rate the topography on which the tea is grown

[1] Steep slope [2] moderate slope [3] Gentle slope [4] other

c) Who has the control over decisions on how the income from tea is spent

[1] Male [2] Spouse/wife [3] Both [4] other

d) From the below options; kindly select one that best describes the way your tea farm is operated

Options	Husband & wife jointly work on same field	Man & wife work on separate fields	Not applicable (spouse is deceased)
Code	=1	=2	=99

e) Please provide information on your **tea farm (s)/fields and yields for 2014/2015** (owned fields only**)**

	Field /Garden	No of bushes	Acres operated	Year Planted	Who planted (1=self; 2= inherited 3= purchased)	Variety planted (1 =clonal; (2 =seedling; 3= both)	Annual yield (Kgs)
I	Tea Field(s) operated jointly by man and wife						
Ii	Field(s) operated by husband [1]						
iii	Field operated by wife						
	Total (all fields/						

f) If the tea farm was inherited; what was the initial farm size before sub division?

No of acres	

[1] If they farm separately use row (ii) for husband and row (iii) for wife

g) In case the household leased some tea fields in the last year 2014/2015, please provide information on area and yields obtained (leased fields only**)**

Field /Garden	No of bushes	Acres operated	Year Planted	Variety planted (1 =clonal; (2 =seedling; 3= both)	Annual yield (Kgs)
Tea Field(s) operated jointly by man and wife					
Field(s) operated by husband					
Field operated by wife					
Total (all fields/ Gardens)					

h) How many times do you pluck your tea farm (s) or fields in a month (plucking rounds)?
Once [1] Twice [2] Thrice [3] four times [4]

6) Fertilizer and Manure application

No	Item	Code/Response
(a)	Did you apply fertilizer on your farm in 2014/2015 (yes=1, No=0)	
(b)	How many times did you apply the fertilizer 2014/2015 (Once =1, twice =2, other =3 (specify)	
(C)	How many 50 kgs bags did you apply on your farm (No of bags)	

7) Labour Allocation

a) What is the source of labour for each of the following tea farming activities

Activity	**Indicator**	**response**
Weeding	1= family; 2 =employed workers ; 3= hired workers on daily rate; 4= others (specify)	
Plucking	1= family; 2 =employed workers ; 3= hired workers on daily rate; 4= plucking rate /kg; 5= others (specify)	
Fertilizer application	1= family; 2 =employed workers ; 3= hired workers on daily rate; 4= others (specify)	
Delivering tea to collection centre	1= family; 2 =employed workers ; 3= hired workers on daily rate; 4= others (specify)	
Others	1= family; 2 =employed workers ; 3= hired workers on daily rate 4= others (specify)	

b) How many man days were allocated to each of the following tea farming activities in the last year (2014/15) (use table below)

Activity				Hired				Family Labour					
	Site	# of male hired workers	# of days	# of female hired workers	# of days	Kshs / man day	Total Kshs	# of males	# of days	# of females	# of days	# of children	# of days
Weeding	1												
	2												
	3												
Total weeding													
Plucking	1												
	2												
	3												
Notes													
Total Plucking													
Fertilizer application	1												
	2												
	3												
Total fertilizer application													
Transport and waiting time tea (BC) (all fields)													
Other, specify	1												
	2												
	3												
Total for others													
Calculated Total													

a) What is the daily wage rate for general farm labour in this area?_____(Ksh per day)

b) For this wage, what is the typical number of hours worked per day?_____(hrs)

c) What is the plucking rate per kg----------

8) Information on tea Marketing Channels

a) Where did you sell your green leaf in 2014/15?
[1] Exclusively KTDA [2] Otherwise

b) Kindly indicate the volume of green leaf that was sold through each of the channels in last year (2014/15)

	Channel through which green leaf was sold	Quantity sold through channel in kgs (last year)	Price/kg (initial)	Bonus /kg (Bonus)	Total
1	KTDA Factory				
2	Private Factory				
3	Middleman				

c) If you sold green leaf through a non KTDA factory, kindly give reasons

……
.....
.....

d) If you only sold green leaf through a KTDA factory, kindly give reasons
.....
.....
.....

e) Now, as compared to 5-8 years ago, is it more convenient or less convenient to sell your tea?

[1 = more convenient; 2 = Less convenient; 3 = the same; 4 = not sure; 99= not applicable]

9) **Access to Information on training received in 2014/2015**

a) Do you have access to extension services?
1= Yes 2= No

b) If yes state the frequency of your interaction with (use table below)

Frequency of interaction (see code above)	(KTDA)	Private factory	Other (specify)
2=weekly			
3=monthly			
4=once in 3 months			
5=once in 6 months			
6=once in the year			
7= other (specify)			

c) Have you ever participated in A farmer field school (FFS)

Response	yes =1	No=2

d) Which other training activity on tea have you participated in recent past (3 years)
[1] Field days [2] Rain forest (RA) [3] others specify

10) Notable tea farming technologies or practices : For each of the following technologies or husbandry practices, please indicate the usage

Technology / husbandry Practice	(a) Do you know this technology? 1=Yes 0=No	(b) Have you ever used this technology in your tea fields 1=Yes 0=No	(c)Year you first used this technology	(d) If you did not use this technology during the 2010/11 season why?
Pruning machine				
Manure				
Pruning height				
3 - 4 plucking rounds /month				
Plucking table				
Spacing				

11) Tea farming constraints

a) did the household experience any problem in accessing any of the following inputs

		Plucking Labour	Weeding labour	Fertilizer
Did you experience problems	1= yes 2= no			
What was the main reason behind inaccessibility	1=scarcity 2=lack of cash 3= delay			
How did that affect your input use	1= did not affect 2= reduced application 3= delayed application			

b) When you think about your tea farming activities; are the following a significant tea farming constraint (Tick the appropriate response)

No		**Strongly agree=5**	**Agree= 4**	**Neutral=3**	**Disagree= 2**	**Strongly disagree=1**
1	Availability of labour					
2	Tea collection					
3	Access to fertilizer					
4	Regulations					
5	Factory management					
6	payment					
7	prices					
8	Accountability					
9	Cost of inputs					
10	Deductions					
11	Road infrastructure					
12	Weather					
13	Tea rejections at BC					
14	Pests and diseases					
15	Other specify					

Section C: Further Household Characteristics

12) Information on household health status in (last season/year)

No	Health status	Code/response
1	During the past 12 months, did anyone in this household suffer from any illness or injury? (1= Yes 2 = No)	
2	For how many days did the affected household members have to stop doing their usual activities due to illness or injury?	
3	How frequently did a members or members of your household visit a medical facility **Codes:** 1= not all, 2= Once in a while, 3=Monthly 4= Regularly	
4	Summary Index on health status (derived from 1-3 above (1= good, 2=average 3= poor)	

13) Information on household food security/ self-sufficiency status in 2014/15

No	Food security status	Code/response
1	In the last 12 months, did the household rely entirely on Staples from own production (1= Yes 0 = No)	
2	If no to 1 above , how many months did the household rely on purchased main food staples (no of months)	

14) **Information on household wealth indicators/conditions**

a) **Housing conditions**; to be observed by enumerators. If they can't tell, or aren't sure, ask the respondent

What is the major construction of the material of the roof? 1=Thatch 2=Iron sheets; 3=Tiles; 4= Other (specify)		What is the major construction material of the external wall? 1 = Mud and poles; 2= Timber; 3= Burnt bricks/stone /blocks with cement; 4= Other (specify)		What is the major material of the floor? 1= Earth; 2 = Cement; 3= tile; 4= Other (specify)	

b) Households assets

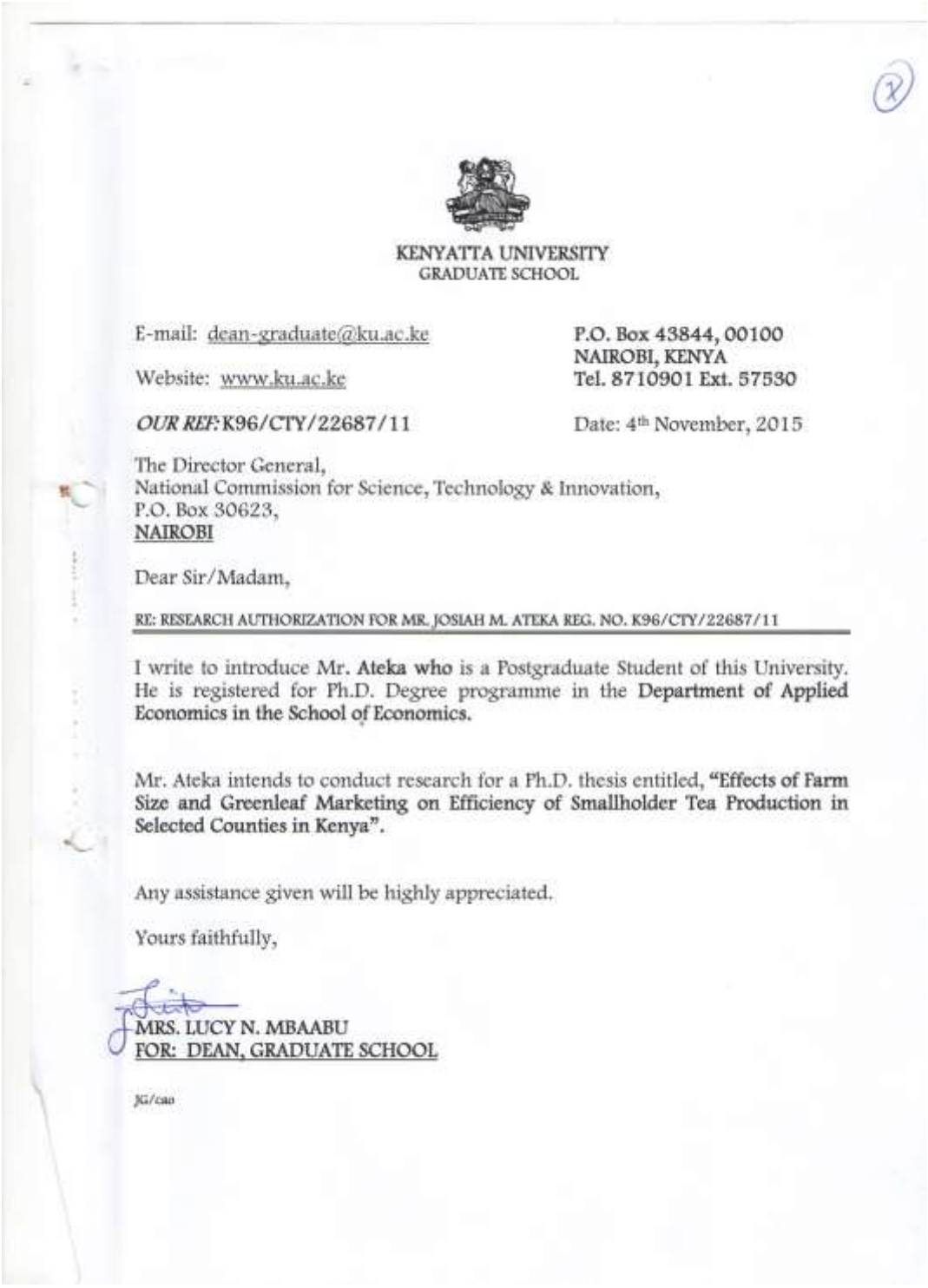
Which of these assets does your household own?	1 = Yes 2 = No	Approximate value	Which of these assets does your household own?	1 = Yes 2 = No	Approximate value
Bicycle			Motorcycle		
Radio			Wheelbarrow		
Mobile phone			cows (indicate No)		
Television			No of chairs		
Motor car			No of chicken		
Others (specify)					

15) Information on Household Expenditure

On the average, how much do you spend on the following food and non-food items per Month? (If respondent does not know how much they consumed, focus on how much they spent on the item).

	Food				Non Food Item			
	Items:	Qty (kg)	Unit price	Exp (ksh)	**Items:**	Qty (kg)	Unit price	Exp (ksh)
1	Maize / maize flour				Rent			
2	Beans				Telephone (airtime, charging)			
3	Rice				Transport			
4	Vegetables				Vehicle			
5	Milk				Electricity			
6	Beverage:				Fuel			
7	Bread				Fire wood			
8	Fruits				Religious costs			
9	Meat				Condolence			
10	Cooking Fats / oil				Medical			
11	Alcohol / cigarette				Shoes			
12					Books			
13					Clothing			
14					School fees			
15					School uniforms			
	Analyst to calculate annual Expenditure				Analyst to calculate annual Expenditure			

7.4 Appendix 4: Research Authorization



KENYATTA UNIVERSITY
GRADUATE SCHOOL

E-mail: dean-graduate@ku.ac.ke
Website: www.ku.ac.ke

P.O. Box 43844, 00100
NAIROBI, KENYA
Tel. 8710901 Ext. 57530

OUR REF: K96/CTY/22687/11

Date: 4th November, 2015

The Director General,
National Commission for Science, Technology & Innovation,
P.O. Box 30623,
NAIROBI

Dear Sir/Madam,

RE: RESEARCH AUTHORIZATION FOR MR. JOSIAH M. ATEKA REG. NO. K96/CTY/22687/11

I write to introduce Mr. **Ateka who** is a Postgraduate Student of this University. He is registered for Ph.D. Degree programme in the **Department of Applied Economics in the School of Economics.**

Mr. Ateka intends to conduct research for a Ph.D. thesis entitled, **"Effects of Farm Size and Greenleaf Marketing on Efficiency of Smallholder Tea Production in Selected Counties in Kenya".**

Any assistance given will be highly appreciated.

Yours faithfully,

MRS. LUCY N. MBAABU
FOR: DEAN, GRADUATE SCHOOL

JG/cao

7.5 Appendix 4: Research Permit