MAIZE PRODUCTION INTENSIFICATION IN KENYA

By

John Otieno Olwande

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Community Sustainability – Doctor of Philosophy

2018

ABSTRACT

MAIZE PRODUCTION INTENSIFICATION IN KENYA

By

John Otieno Olwande

Soil infertility is one of the major problems contributing to low and stagnated agricultural productivity in sub-Saharan Africa. In Kenya, this problem is manifested in maize where yields have remained low and stagnated over time despite increased use of inorganic fertilizers and improved seed varieties. More effective alternatives and/or complimentary actions to address this problem thus remains germane. This dissertation contributes to that endeavor through two broad objectives: to generate evidence that can support decisions to address the problem of low agricultural productivity in general and of maize in particular in Kenya; and to contribute to the body of knowledge about agricultural intensification in sub-Saharan Africa.

The first essay uses household panel survey data from rural Kenya covering a period of 13 years (1997 – 2010) to examine trends and patterns in land and labor productivity of maize, measured as net returns to land and to family labor. Results show declining landholdings and farm sizes but maize occupied over one-half of cultivated land. Land productivity declined by 42% for households with at least 10 acres and by 33% in the most important maize producing regions. Labor productivity increased in areas with smaller landholdings and higher population density because of increase in land productivity, and declined or only marginally increased in areas with larger landholdings and lower population density because of a decline in land productivity. These results demonstrate that increasing maize production and returns to family labor in Kenya will rely on improving yields especially in the major maize growing areas where this has declined.

The second essay uses data on maize production in five major maize growing counties in Kenya to compare maize farmers' perceived soil fertility to measured soil fertility. It also investigates the influence of farmers' perceptions of soil fertility on their adoption (use) of soil fertility management practices. Results show little agreement between farmers' perceived and measured soil fertility, and farmers mostly judge the fertility status of soil by crop performance. Farmers apply management practices that may not match the fertility needs of soil on their plots, exemplified by the persistent application of an acidifying fertilizer (diammonium phosphate (DAP)) and low application of organic soil amendments even on plots with soils that are acidic and deficient in organic carbon. Farmers on average are more likely to apply inorganic fertilizer to plots they perceive to be infertile, and they treat manure or compost and inorganic fertilizers as serving substitute roles in soil fertility. These results suggest policy and extension information gaps regarding soil fertility management.

The third essay uses the same dataset as in the second essay together with rainfall data to estimate technical efficiency of maize farmers and the effect of farmers' soil fertility perception on technical efficiency. It also demonstrates the importance of including environmental production conditions and agronomic practices in agricultural productivity and efficiency analysis. Average technical efficiency level is 0.75 and 0.70, respectively, with and without environmental variables and agronomic practices in the model, indicating that scope for increasing maize yield through better management of inputs exits and that omission of environmental variables and agronomic practices underestimates technical efficiency. Farmers' perception of soil fertility and the consistency of their perception with measured soil fertility both have significant effects on technical efficiency, underscoring the importance of information that can enhance farmers' accurate understanding about soil fertility conditions on their farms to help them make better production decisions.

Copyright by
JOHN OTIENO OLWANDE
2018

For my family
Salome, Chamola, Atara, Hawi
For my mother
Risper Olwande
In memory of my late father
Olwande Diang'a
For my step-brother
Eric Olwande

ACKNOWLEDGEMENTS

I express my sincere gratitude to Dr. John Kerr, my major advisor and dissertation committee chair. The completion of this dissertation and my entire PhD program is because of his invaluable guidance, mentorship and motivation. In addition to my major advisor, I also thank my advisory and dissertation committee members, Dr. Steven Gray, Dr. Thomas Jayne and Dr. Robert Myers, for their guidance in my research and writing process and during the entire program. Their insightful suggestions and constructive criticisms greatly shaped my approach to research and writing.

The completion of my PhD program would have been a dream had it not been for the sacrificial support I received from Dr. John Kerr and Dr. Thomas Jayne. I particularly am deeply grateful to Dr. Jayne for encouraging and supporting me to join the PhD program at Michigan State University (MSU), to Dr. Kerr for accepting me to be his student and to both for generously funding my studies. I also thank Dr. Melinda Smale, Dr. Milu Muyanga and Dr. Miltone Ayieko for their motivation and encouragement before and after I joined the program.

My special gratitude goes to the Department of Community Sustainability (CSUS) for offering me the George and Nancy Axinn International Scholarship award to support my studies. I also thank the Food Security Group at the Agricultural, Food and Resource Economics (AFRE) department and Tegemeo Institute of Egertion University, Kenya, for providing me the data for dissertation research. I particularly thank Dr. Thomas Jayne and Dr. Milu Muyanga of AFRE and the Director of Tegemeo, Dr. Miltone Ayieko, for facilitating my access to the data.

My sincere thanks to fellow graduate students with whom we shared moments of joy and frustrations and breakthroughs and dead ends during the program. I appreciate the support I received from, among others, Socheat Keo, Mayuko Kondo, Awa Sanou, Aryal Giri, Nahid Sattar, Yu Jina, Soo Jeong, KJ Lee, Obafemi Elegbede, Arfang Dabo, Godwin Debrah, Juma Mmongoyo and Chewe Nkonde. My special thanks to Brendan O'Neill for helping me understand the different soil quality parameters.

I extend thanks to the families of Dr. Miltone Ayieko and Dr. Milu Muyanga who welcomed me to East Lansing and made me part of them throughout my program. I also express my special thanks to the family of Drs. Rex and Vangie Alociljia and the entire International Life Group for their emotional and spiritual support through the weekly bible study meetings and constant prayers.

My special and sincere thanks go to my mother, Risper Olwande and late father, Olwande Diang'a. Through their relentless toil tilling the land, they provided the much they could for me and my siblings. Although they did not have the means to provide decent education for me and my siblings, their words of encouragement and messages of hope to us motivated me to seize available opportunities to acquire education and always aim higher.

I am deeply grateful to my uncle Humphrey Omwomo and step-brother Eric Olwande for the sacrifice they made to pay for my school fees. I particularly am indebted to Eric for his relentless effort and sacrifice in paying for my schooling when I had lost hope in ever completing high school. His sacrificial support to my education will remain a permanent mark in my professional life. May God richly reward him! I also thank my brothers and sisters who always encouraged me in many ways.

My family has been instrumental to my successful completion of the PhD program. I am indebted to my wife Salome Olwande and our son and daughters, Chamola Olwande, Atara Olwande and Hawi Olwande. The sacrifice they made to allow me live and study thousands of miles away from them is invaluable. The love, patience and emotional support they provided and continue to provide me are simply overwhelming. May God bless them!

Finally, I thank the Almighty God for His favor throughout this journey. May all glory remain with Him!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Agriculture remains the major livelihood source for the majority of the population in sub-Saharan Africa and is essential for food security, poverty reduction and economic growth (The World Bank, 2007). Despite this fundamental role, agricultural productivity in sub-Saharan Africa is generally low and its growth has stagnated (Haggblade & Hazell 2010; Otsuka & Larson, 2013). Among the factors responsible for the poor performance of Africa's agriculture are land degradation, small and declining farm sizes because of population growth, reliance on rain-fed farming, low use of external inputs in production and poor market infrastructure (Barbier & Hochard, 2014; The Montpellier Panel, 2013; Haggblade & Hazell, 2010; Otsuka & Larson, 2013). Improving food security, reducing poverty and realizing faster economic growth in this part of the world requires accelerated and sustainable agricultural productivity growth (The World Bank, 2007; Diao et al, 2007).

The concern about low and stagnated agricultural productivity in sub-Saharan Africa applies to Kenya as well. This is most evident in maize, the most important staple grain and which is widely grown, where aggregate yield has exhibited a declining trend over the past quarter century. Surprisingly, the declining trend in yield has occurred despite impressive growth in inorganic fertilizer and improved seed use and remarkable public investments in rural infrastructure over the years (Ariga & Jayne, 2009; Chamberlin, 2013; Smale & Olwande, 2014). Maize consumption demand, on the other hand, has been on an upward trend due to sustained population growth, resulting in a widening gap between consumption demand and production and increasing import bill (Kirimi et al., 2011).

Because of its crucial role as the main staple food for a large section of the population, investment in achieving sustainable productivity growth in maize remains a key goal in the government's policy and development strategies (Government of Kenya, 2010a, Government of Kenya, 2010b) and development organizations' support agendas. The government, international development agencies and non-governmental organizations (NGOs) have continued to implement a range of interventions, spanning the farm and input and output markets in efforts to boost maize production. Notable among the interventions are the state's fertilizer subsidy programs and producer price support, alongside a strong establishment for seed research and development and investments in rural infrastructure.

Clearly, we can expect these efforts to have enhanced maize production through faster growth in yield in Kenya. The lack of yield growth in maize despite the largest price support, input subsidies, and seed research efforts is puzzling. The scenario indicates existence of factors beyond what the current efforts are delivering that may be playing a significant role in stymied maize yield, and demonstrates the need for more effective alternatives and/or complimentary actions that can nurture investments in greater production intensification in a sustainable way.

Smale & Olwande (2014) note that several explanations for the dismal performance in maize yield have been advanced, including degraded soils associated with land pressures arising from population density, reduced fallows and nutrient mining, resulting in poor crop response to fertilizer use. Some of the explanations, however, have not been corroborated by evidence from rigorous empirical investigation to understand the nature of the effects of these factors on maize yield and the extent to which they may be contributing to the observed low and stagnated productivity. Therefore, the search for strategies to address the problem of low and stagnated agricultural productivity in Kenya remains germane.

This dissertation contributes to addressing the aforementioned need. It pursues two overarching objectives. The first is to generate empirical evidence to support decisions that can address the problem of low agricultural productivity generally and of maize in particular in Kenya. The second is to contribute to the body of knowledge about agricultural intensification in sub-Saharan Africa, given the increasingly recognized fact by African governments that sustainable farm productivity growth is key to achieving sustainable economic growth and improving the livelihoods of hundreds of millions of their constituents. These overarching objectives are pursued through three separate but linked essays.

The first essay (Chapter 1) is entitled "Trends and Patterns in Land and Labor Productivity in Kenya". In addition to low and stagnating productivity, farm sizes in Kenya have been shrinking over time with rising rural population densities and sub-division. Moreover, the long-term low and stagnated agricultural productivity growth may have depressed the generation of dynamic farm-nonfarm growth multipliers and shifts in the composition of the labor force that have been the foundation of economic transformation in other regions of the world (Mellor, 1976; Johnston & Mellor, 1961; Lipton, 2005). This essay seeks to determine whether Kenya's agricultural sector is changing in ways that are promoting or retarding farm labor productivity, and to understand the association between farm labor productivity and population density, land scarcity and market access. Farm labor productivity is defined in terms of net returns to family labor, a measure that is deemed most appropriate for profitability of farming for agricultural households.

The essay is focused on maize production and uses a nationwide five-wave (1997, 2000, 2004, 2007 and 2010) panel survey dataset on a sample of about 1500 agricultural households in Kenya. Descriptive and bivariate analyses are conducted to examine trends and patterns in the size of landholding, area planted and land allocation to maize using data for all the survey years, and in land and labor productivity using data for 2004, 2007 and 2010 for which labor data was available.

Six key results emerge. First, average landholding declined by over 12% while average area planted declined by 11%. Maize occupied over 50% of total area planted. The average area of plots with maize for a household declined by 11% over time. Second, long-term labor productivity (returns to family labor) increased by 38% overall, contributed by positive changes in land-labor ratio of 16% and land productivity of 21%. The increase in land-labor ratio was because of a decrease in family labor use rate on maize plots. Third, there was an inverse relationship between landholding and land productivity and landholding and family labor use rate. Average labor productivity increased for all households. However, the increase in labor productivity for those with 10 or above acres was entirely because of positive change in land-labor ratio rather than in land productivity, which declined by 42%. Fourth, average land productivity and family labor use rate had direct relationships with population density while labor productivity was inversely related with population density. Fifth, most remote households compared to their least remote counterparts had lower land productivity, higher family labor use rate and lower labor productivity, on average. Lastly, there was a remarkably small increase in labor productivity (0.7%) in the High Potential Maize zone (HPMZ), Kenya's most important maize producing region, because of a 33% decline in land productivity. These results demonstrate that increasing maize production through land expansion is infeasible. Increasing production and labor productivity will have to rely on yield (land productivity) growth. This is especially urgent for the major maize growing areas where yield has declined yet they are the main suppliers of maize in the domestic market.

The second essay (Chapter 2) is entitled "Farmers' Perceptions and Adoption of Soil Fertility Improvement Practices". The evidence of increased adoption of fertilizer and improved maize seed over time without significant growth in maize yield implies that continued use of fertilizers and improved maize varieties on their own cannot achieve the needed growth in yield. There is need for sustainable management practices that add organic matter and ameliorate soil acidity to restore fertility and subsequently increase crop response to external inputs (Lal, 2006; Chivenge et al., 2011; Kunhikrishnan et al., 2016; McCauley et al., 2017). We can expect that farmers' perceptions about the fertility conditions of their soils bear on their decisions about adoption of technologies and agronomic practices to improve soil fertility. Few quantitative studies exist about the correspondence of farmers' perceptions about the fertility conditions of their soils and measured soil fertility, and how those perceptions influence adoption (or use) of soil fertility management technologies and agronomic practices. This essay compares maize farmers' perceptions about the fertility of their soils to soil fertility as measured by results from scientific test of soil chemical properties. It also investigates farmers' adoption (use) of soil fertility management practices, with a focus on the influence of farmers' perception about soil fertility. The essay makes a major contribution of incorporating farmers' perception as a factor in quantitative assessment of their soil fertility management decisions.

The essay uses two cross-sectional household- and plot-level survey data on maize production spanning five counties located in the major maize growing areas of Kenya. The first survey was conducted in 2014 from a sample of 650 farm households spread in five counties, while the second survey was conducted in 2016 on the same households as in the first survey, but the number reduced to 623 due to attrition. The maize plots that were cultivated in the 2014 survey were not necessarily targeted in the 2016 survey hence the non-panel nature of the plot-level data. Plot-level data include laboratory-tested soil physical and chemical properties.

The relationship between farmers' perceived and measured soil fertility and adoption of soil fertility management practices are analyzed using three approaches. First, a chi-square test is used to examine statistical independence between farmers' perceived and measured soil fertility. Second, interrater agreement technique (Fleiss et al. 2003) is used to estimate the degree of agreement between farmers' perception of soil fertility and measured soil fertility. Third, a probit model is estimated to examine the relationship between farmers' perceived and measured soil fertility in the presence of other relevant factors. Lastly, a multivariate probit model is estimated to identify the effect of farmers' perception about soil fertility on their adoption of soil fertility management practices.

Results show that there is little correspondence between farmers' perceived and measured soil fertility, and farmers mostly rely on crop performance to judge the fertility status of their soil. Soil testing to understand the fertility condition of soil was a rare practice. These results imply that farmers' management practices may not be compatible with the fertility needs of their soils, as exemplified by the persistent application of an acidifying fertilizer (DAP) and low application of organic soil amendments such as manure and compost even on soils that are acidic and low in organic carbon, potentially worsening the problem. Farmers' soil fertility perception has a strong relationship with the decision to apply inorganic fertilizer, but not other soil fertility management practices such as applying manure or compost; the likelihood of applying inorganic fertilizer to a plot increases when a farmer perceives a plot to be infertile. Farmers also appear to treat manure/compost and inorganic fertilizers as substitutes. These results reflect gaps in policy and extension, exemplified by the state's aggressive promotion of use of inorganic fertilizers, especially DAP, without accompanying concerted efforts to promote use of organic soil amendments.

The third essay (Chapter 3) is entitled "Technical Efficiency and Soil Fertility and Agronomic Practices in Maize Production in Kenya". Lack of land for agricultural expansion in Kenya, and indeed Africa (Chamberlin, Jayne, & Headey, 2014), implies that increasing agricultural productivity is the feasible option that can spur broad-based economic growth, help reduce poverty and enhance food security. Increasing agricultural productivity can be achieved through technological change, increased efficiency in use of existing technology and productive resources or both. This essay estimates technical efficiency level of maize farmers to determine potential productivity gains possible through better management of production resources. It identifies factors that affect variations in technical efficiency across farms, with a focus on the effect of farmers' soil fertility perception, and demonstrates the importance of including environmental production conditions and agronomic practices in agricultural productivity and efficiency analysis. Much of the literature on agricultural productivity and efficiency analysis have often ignored farmers' perceptions of soil fertility, environmental production conditions and agronomic practices, in part due to data limitations, yet these factors can be expected to condition input choice and use decisions and subsequently productivity and efficiency.

The data used in this essay is the same as in essay 2. In addition, the essay uses rainfall data extracted from the Climate Hazards group Infrared Precipitation with Stations (CHIRPS) dataset (Funk et al., 2015). The stochastic production frontier (SF) approach due to Meeusen & van Den Broeck (1977) and Aigner et al (1977) is used to estimate technical efficiency of maize farmers, with and without controlling for environmental conditions (soil fertility conditions and rainfall) and agronomic practices, and identify factors responsible for variation in technical efficiency across farms.

Three key results have emerged. First, average level of technical efficiency is 0.75 and 0.70, respectively, with and without controlling for environmental variables and agronomic practices. Inefficiency-induced foregone maize yield is 0.34 tonnes/acre when environmental variables and agronomic practices are controlled for and 0.45 tonnes/acre without controlling for them. The forgone output is substantial considering that the reported yield is only about 1.3 tonnes/acre. Secondly, omission of environmental production conditions and agronomic practices from the stochastic frontier model results in understated technical efficiency levels. Lastly, farmers that view their plots as fertile have, on average, 19% lower technical inefficiency than those who view their plots as infertile. In addition, those whose perceptions about the fertility status of their plots are consistent with measured soil fertility have 2.9% lower technical inefficiency than those whose perception is inconsistent with measured soil fertility, on average. These results indicate that scope for increasing maize yield through better management of inputs exists and underscore the necessity for farmers' accurate understanding about soil fertility conditions on their farms to help them make better production decisions. Failure to account for environmental production conditions and agronomic practices in productivity and efficiency analysis may yield inaccurate results.

REFERENCES

REFERENCES

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

Ariga, J., & Jayne, T. S. (2009). *Private Sector Responses to Public Investments and Policy Reforms The Case of Fertilizer and Maize Market Development in Kenya* (IFPRI Discussion Paper No. 921).

Barbier, E. B., & Hochard, J. P. (2014). *Land degradation, less favoured lands and the rural poor: A spatial and economic analysis*.

Chamberlin, J. (2013). *Market access and smallholder development in Kenya and Zambia*. Michigan State University.

Chamberlin, J., Jayne, T. S., & Headey, D. (2014). Scarcity amidst abundance? Reassessing the potential for cropland expansion in Africa. *Food Policy*, *48*(2014), 51–65. https://doi.org/10.1016/j.foodpol.2014.05.002

Chivenge, P., Vanlauwe, B., & Six, J. (2011). Does the combined application of organic and mineral nutrient sources influence maize productivity ? A meta-analysis. *Plant Soil*, *342*, 1–30.

Diao, X., Hazell, P., Resnick, D., & Thurlow, J. (2007). *The Role of agriculture in development: Implications for sub-Saharan Africa*. Washington, D.C., USA. https://doi.org/153

Fleiss, J. L., Levine, B., & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., … Michaelsen, J. (2015). Climate Harzads Group. https://doi.org/10.1038/sdata.2015.66

Haggblade, S., & Hazell, P. B. R. (Eds.). (2010). *Successes in African agriculture: Lessons for the future*. Baltimore: The Johns Hopkins University Press.

Johnston, B. F., & Mellor, J. W. (1961). The role of agriculture in economic development. *The American Economic Review*, *51*(4), 566–593.

Kirimi, L., Sitko, N., Jayne, T. S., Karin, F., Muyanga, M., Sheahan, M., … Bor, G. (2011). *A farm gate-to-consumer value chain analysis of Kenya's maize marketing system* (Working Paper No. 44). Nairobi.

Kunhikrishnan, A., Thangarajan, R., Bolan, N. S., Xu, Y., Mandal, S., Gleeson, D. B., … Naidu, R. (2016). Functional Relationships of Soil Acidification, Liming, and Greenhouse Gas Flux. *Advances in Agronomy*, *139*, 1–71. https://doi.org/10.1016/bs.agron.2016.05.001

Lal, R. (2006). Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, *209*(August 2005), 197–209.

Lipton, M. (2005). *The family farm in a globalizing world: The role of crop science in alleviating poverty* (2020 Discussion Paper No. 40). Washington, DC.

McCauley, A., Jones, C., & Olson-Rutz, K. (2017). Soil pH and organic matter. *Nutrient Management*, (8), 1–16.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error, *18*(2), 435–444.

Otsuka, K., & Larson, D. F. (Eds.). (2013). *An African Green Revolution: Finding ways to boost productivity on small farms*. Dordrecht Heidelberg New York London: Springer. https://doi.org/10.1007/978-94-007-5760-8

Smale, M., & Olwande, J. (2014). Demand for maize hybrids and hybrid change on smallholder farms in Kenya. *Agricultural Economics (United Kingdom)*, *45*(4).

The Montpellier Panel. (2013). *Sustainable intensification: A new paradigm for African agriculture*. London.

The World Bank. (2007). *World development report 2008: Agriculture for development*. Washington, DC: The World Bank.

CHAPTER 1: TRENDS AND PATTERNS IN LAND AND LABOUR PRODUCTIVITY OF MAIZE IN KENYA

1.1 Introduction

Despite recent trends showing a shift in the labor force out of farming, the agricultural sector remains the single most important source of livelihoods for most people in Sub-Saharan Africa (Yeboah & Jayne, 2016). However, agricultural performance has been largely disappointing (Haggblade & Hazell, 2010). Land and labor productivity levels have remained low and stagnated and per capita food production has been declining (Otsuka & Larson, 2013). African agriculture is characterized by continued reliance on rain-fed and mostly low-input production; farms that are small and declining in size over time due to population growth; a trend toward increased land degradation (Barbier & Hochard, 2014; The Montpellier Panel, 2013); and marketing infrastructure that, while improving, remains underdeveloped and imposes high costs on participants in the agri-food system (Haggblade & Hazell, 2010; Otsuka & Larson, 2013).

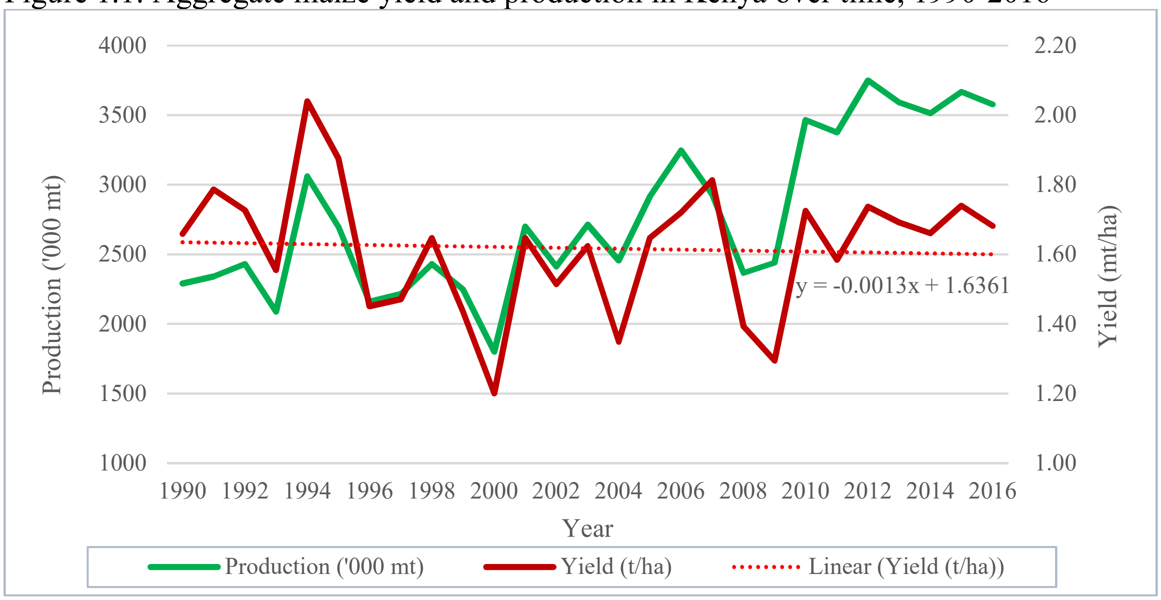
The underperformance of African agriculture and the recognition of the central role of agricultural productivity growth in development in the region has triggered a renewed interest in agriculture by African governments and in international development circles (Otsuka & Larson, 2013; The World Bank, 2008). While there is a broad consensus that an agricultural productivity revolution is necessary to foster economic growth in sub-Saharan Africa, Otsuka & Larson (2013) note that there is no consensus on how to realize that revolution. The authors suggest that debate on strategies to adopt to promote general agricultural growth revolves around whether to focus on small or large farms, to prioritize food staples or high value products, to promote production practices that rely on fertilizers and modern seed varieties, and the extent that governments should get involved in markets. In their efforts to promote agricultural productivity, over the last decade many African governments have revived input subsidy programs to encourage the use of chemical fertilizers and modern seed varieties on food staples, and/or operated directly in food markets to raise producer prices for farmers (Jayne & Rashid, 2013).

This description of the region's general trends in agricultural productivity applies in most respects to Kenya. Agricultural productivity is low and has generally stagnated over the long term. Specifically, yield of maize, the most important staple grain which is grown by virtually every agricultural household in the country, has shown a declining trend in the long term (Figure 1.1). The long-term picture is a result of sustained decline in yield witnessed in the period 1990-2003, which has not been offset by the slow rate of increase in yield observed between 2004 and 2015, and levels of which have remained lower than in early 1990s (Figure 1.2). The yield increase observed after 2004 is attributable to a range of efforts at revitalization of agriculture as part of the strategy to revive Kenya's economy after decades of persistent decline. The efforts focused mainly on reviving collapsed and dysfunctional agricultural institutions that offer services to farmers, policy and regulatory reforms and investments that facilitate farmers' access to input and output markets (Government of Kenya 2010).

In addition to low and stagnated productivity, farm sizes have been shrinking over time with rural population densities and sub-division, with most affected areas being those with high agro-ecological potential and where population densities are highest (Muyanga & Jayne, 2014). The pressure on farm land because of population increase is one of the factors responsible for soil degradation from nutrient mining, which is one of the main reasons for low and stagnated agricultural yields (Government of Kenya, 2014; Tittonell et al., 2008; Marenya & Barrett, 2007).

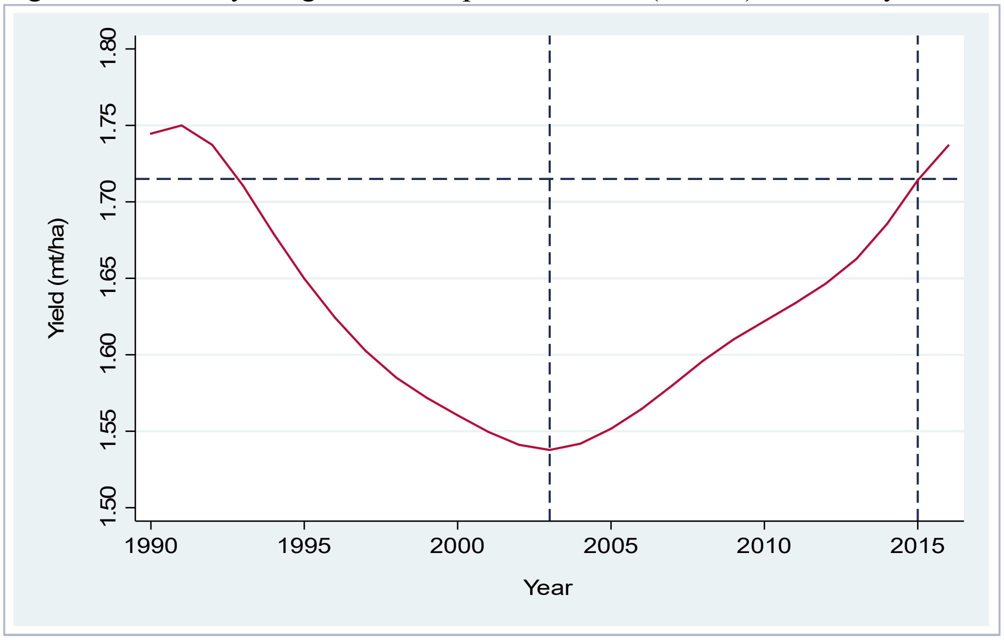
Rapidly rising population combined with stagnant agricultural productivity growth and limited potential for cropland expansion is making Kenya increasingly dependent on food imports for its national food security (FAOSTAT, 2014). Moreover, the country's record of relatively low agricultural productivity growth over the past several decades has depressed the generation of dynamic farm-nonfarm growth multipliers and shifts in the composition of the labor force that have been the foundation of economic transformation in other regions of the world (Mellor, 1976; Johnston & Mellor, 1961; Lipton, 2005).

Figure 1.1: Aggregate maize yield and production in Kenya over time, 1990-2016



Source: Author's compilation using data from the *Ministry of Agriculture, Livestock and Fisheries (MoALF)*
Note: Yield is computed from production and area estimates by the MoALF. Area data includes those of mono-cropped and inter-cropped maize plots. So, yield includes all plots with maize on them.

Figure 1.2: Locally weighted scatterplot smoother (lowess) of maize yield over time



Source: Author's computation using data from *Ministry of Agriculture, Livestock and Fisheries (MoALF)*

The broad objectives of this study are to: (1) determine whether Kenya's agricultural sector is changing in ways that are promoting or retarding farm labor productivity; and (2) understand the association between farm labor productivity and population density, land scarcity and market access. We define farm labor productivity in terms of net returns to family labor, a measure that we deem most appropriate for profitability of farming for agricultural households. We attempt to address the following questions in pursuit of these objectives:

- How has land area planted to maize changed over time?
- How has labor productivity of maize changed over time? What is the pattern in labor productivity across different segments of population density, market access and landholding?

- How are increased population density, smaller farm sizes and market access associated with labor productivity of maize? What does this mean for agricultural intensification in Kenya?

1.2 Conceptual framework

Boserup (1965) posited that rising population density exerts pressure on access to land, inducing farmers to adopt more intensive systems of land use. Farmers move away from extensive to permanent cultivation systems through a gradual transition from long fallow periods, involving shifting cultivation, to multiple cropping of land, where multiple successive crops are planted on the same area of land every cropping year. This change in land use system is accompanied by more intensive use of inputs, such as labor, compost, manure, improved seeds, inorganic fertilizers and irrigation, to increase production.

The link between population density and labor productivity, defined as net returns to family labor, can be summarized through the following identity:

$$y = \frac{Y}{L} \equiv \frac{Y}{A} \times \frac{A}{L} \qquad (1.1),$$

where $Y$is net revenue from output from a plot, L is family labour days used in production on the plot, and A is the area of the plot. The ratio $\frac{Y}{L}$ is labour productivity. It is identically a product of land productivity (yield), $\frac{Y}{L}$, and land-labour ratio, $\frac{A}{L}$, which can be considered as efficiency of family labour (Auer, 1966). Logarithmic transformation of (1.1) and differencing results in (1.2), which shows that family labor productivity growth is the sum of growth in yield and in land-family labor ratio:

$$\Delta \ln y = \Delta \ln\left(\frac{Y}{L}\right) \equiv \Delta \ln\left(\frac{Y}{A}\right) + \Delta \ln\left(\frac{A}{L}\right) \qquad (1.2).$$

Alternatively, we can express the identity in (1.2) as in (1.3):

$$\Delta \ln y = \Delta \ln\left(\frac{Y}{L}\right) \equiv \Delta \ln\left(\frac{Y}{A}\right) + \Delta \ln(A) - \Delta \ln(L) \qquad (1.3).$$

It is straightforward from (1.2) that when land expansion is infeasible, faster growth in land productivity is imperative to improving labor productivity. If population is rising and land sizes are shrinking, and if labor use is increasing in a manner that makes the land-labor ratio $\frac{A}{L}$ to decline faster than growth in land productivity, we would expect labour productivity to decline. This would result in undesirable effects on the economic welfare of agricultural households and can lead to loss of incentives for households to produce maize. However, increase in population might not necessarily translate to increase in family labor use on maize. Households may allocate more of their labor to other farm and non-farm enterprises while maintaining labor use on maize at the same level or even reducing it altogether. If labor use in maize declines more than area under maize declines, then the land-labor ratio will increase. If the increase in land-labor ratio were sufficiently large to offset any decline in land productivity, then we would expect to see a rise in labor productivity.

Therefore, on one hand we would expect a positive association between population density and labor productivity, driven by increased land productivity because of increased input use. On the other hand, an inverse relationship between population density and labor productivity would be possible and would mean that the increase in land productivity that would be expected from increased labor supply and capital inputs is not sufficiently large to compensate for the cost of increased use of labor and the capital inputs. This would suggest an economically unsustainable path for agricultural intensification.

We would expect a positive association between labor productivity and improved access to markets. Input and output market access influences output supply and input demand (Pingali et al, 1987) and availability of relevant market information and transaction costs are key determinants of market access. High population density facilitates the flow of such information and reduces transaction costs (Chamberlin, 2013; Ricker-Gilbert et al, 2014). Thus, market access has a direct relationship with labor productivity.

The relationship between population density and farm sizes is well grounded on the Boserup (1965) hypothesis. We would expect area planted to maize to decline over time as population increases. Similarly, we would expect the share of area planted to maize in total area planted to decline over time, as households are likely to gradually diversify or shift to farm enterprises with higher value, such as horticulture and dairy, in efforts to maximize output per unit of land. However, given the importance of maize in household food supply, we would not expect the share of maize area in total area planted to decline significantly. We would expect these changes to be larger in the least densely populated villages than in the most densely populated villages because in the most densely populated villages scope for further reduction of plot sizes and diversification may be reaching a limit since landholdings have already become very small on average.

1.3 Data source and analysis

We use a nationwide five-wave panel survey dataset covering 13 years on a sample of 1500 agricultural households in Kenya, collected by Egerton University's Tegemeo Institute of Agricultural Policy and Development. The data pertains to the 1996/97, 1999/00, 2003/04, 2006/07 and 2009/10 cropping years, hereafter referred to as 1997, 2000, 2004, 2007 and 2010, the years in which the data were collected. While not considered nationally representative, the sample covers all eight of Kenya's important agro ecological zones (AEZ) spread over 24 districts as defined in 1997. Argwings-Kodhek et al (1999) explain the sampling process in detail. The data contains a range of information about household farm and non-farm activities, including detailed information on cropping, livestock keeping and off-farm earning activities.

First, we examine trends and patterns in the size of landholding, area planted and land allocation to maize using data for all the survey years (1997-2010). The pooled sample has 6,977 observations, distributed across the years as shown in the second column of Table 1a. Because the sample contains only farm households, there are no landless households included in the sample. Landholding is self-reported household owned total land area. Total area planted by a household is calculated as the sum of area of individual plots planted in the main season. The short season is not included to avoid double counting, since plots planted in the main season are often the very ones planted in the short season in two-season areas. Likewise, only main season production is counted as output. Area under maize is measured as the sum of the size of each plot that contains maize. Most plots containing maize in Kenya are inter-cropped, and hence we should interpret the area measure as the size of the plots on which maize was planted.

We conduct a descriptive analysis of trends and patterns in labor productivity on plots with maize on them, as well as decomposed components and change over time. This analysis uses data for the 2004, 2007 and 2010 survey years, which contain detailed information on production, input use, and hired and family labor used on households' largest maize plot in the main season. On average, the share of the largest maize plot in households' total planted maize area in the main season was 89% over the three years, and increased from 86% in 2004 to 91% in 2010. Therefore, we are confident that this analysis captures the vast majority of maize produced by this nationwide sample of farm households.

There are 3977 largest maize plots (observations) in the dataset. We exclude from the analysis the following: observations that do not have family labor use on the largest maize plot, because our interest is in returns to family labor; observations that had less than 20% share of maize in the total value of crops produced on the largest plot with maize; observations with reported area of less than 0.2 acres for the largest plot with maize, because of concerns about potentially large measurement error in computing productivity; and observations that had negative net value of crop production. The final number of observations used in the analysis is 3256, distributed as follows: 1153 in 2004; 1098 in 2007; and 1005 in 2010.

We measure labor productivity in terms of net returns to family labor-day, where a day is in terms of 6-hour farm-work day[1]. Because intercropping of maize with other crops is a common practice (2910 of the largest maize plots had more than one crop, with 66% of these having one or two other crops besides maize), we include the value of these crops in computing productivity. Net returns are computed as total gross value of all crops produced on the largest plot with maize net of the cost of variable inputs (land preparation, seeds, fertilizers, and hired labor) used on the plot. District median crop prices are used to value the crops. The cost of fertilizer is similarly computed for fertilizer applied on the largest plot with maize. For land preparation, seeds and hired labor, the actual expenditure incurred on these inputs on the largest plot with maize are used. Because more than one year of data is used, net returns are adjusted for inflation using the well-known

[1] 6 hours is the average of the number of hours of farm work in a day reported in the data.

Fisher's ideal price index, with 2004 quantities and district median prices for the crops (maize and those intercropped with it) used as the base.

Throughout, we disaggregate the analysis by quintiles of village population density, quintiles of distance to tarmac (paved) road (an indicator of market access), landholding categories and agro-ecological zone. Village population density, an indicator of land pressure, is measured as village landholding, in square kilometers, per capita. It is computed as the ratio of the total number of residents to the sum of landholdings of the sample households in a village. For trends and patterns in labor productivity, we also separate between monocrop versus intercrop maize cropping systems, gender of household head, use of mechanized land preparation and use of hired labor in maize production.

1.4 Results and discussion

We first present trends and patterns in household landholdings, planted land size and land allocation to plots with maize. Trends and patterns in labor productivity and its components are discussed next. Change in labor productivity over time is decomposed into its components and results discussed.

1.4.1 Patterns and trends in landholding and land allocation to maize production

Trends in mean landholding, area planted, and planted area allocated to plots with maize is reported in Table 1.1. Over time, household average landholding reduced by 0.8 acres, representing a decline by over 12% in 13 years. The effect of reduced average landholding is reflected in the reduction in area planted by 11%, from an average of 3.5 acres in 1997 to 3.1 acres in 2010.

Smallholder farmers in Kenya are highly diversified in their production. The crops produced include cereals such as maize, sorghum, millet, beans; roots and tubers such as potatoes and cassava; industrial crops such as tea, coffee and sugarcane; and a range of fruits and vegetables. Mixed farming involving crops and livestock is also a common practice. Although agro-ecological conditions determine the specific crops and livestock species that can be profitably produced in each locality, maize production is supported to various extents in virtually all agro-ecological zones of Kenya. Maize is the single most important crop in terms of planted area allocation, with plots with maize occupying over 50% of the total area planted, on average (Table 1.1). Over time, the average area of plots with maize for a household declined by 11%, commensurate with the decline in average total area planted by a household.

The trends and patterns in Table 1.1 have two important implications. First, the shrinking landholdings and the subsequent decline in area planted among smallholder farmers support evidence presented in Jayne et al (2014) that most farms in Kenya, and indeed the region, are declining in size as rural populations continue to rise with little or no potential for land expansion for agriculture. Therefore, this puts the onus on strategies to achieve agricultural growth through agricultural intensification and productivity growth on existing farmland. Secondly, maize remains fundamental to Kenyan agriculture. Even with a small decline in the proportion of total planted area under maize, still over 50% of planted land in Kenya is devoted to maize and hence one of the most effective entry points for accelerating agricultural productivity growth in Kenya is to raise the yields and profitability of maize.

The declining landholding and farm sizes is largely a result of increasing population. We would expect decline in landholding and farm sizes to be larger in least densely populated villages than in most densely populated villages because in most densely populated villages landholdings and farm sizes may have already become so small that land cannot be feasibly subdivided; in such areas, we would expect high rates of migration of youth out of the area and/or shifts in the local labor force from farm to non-farm activities. Indeed, our data shows that outmigration rate of youth is positively correlated with population density (see Figure A1.1 in the Appendix). The analysis in Table 1.2 shows that overall, average landholdings and area planted, both total and allocated to plots with maize were significantly larger in the least densely populated villages than in the most densely populated villages. The proportion of area of plots with maize to total area planted was also significantly larger in the least densely populated villages.

Table 1.1: Levels of and changes in mean landholding, total area planted, and area allocated to plots with maize 1997-2010

Year	N	Total landholding (acres)	Total area planted (acres)	Area planted with maize (acres)	Proportion of maize area in total area planted
1997	1499	6.1	3.5	1.8	0.59
2000	1441	6.0	4.1	2.1	0.57
2004	1396	6.3	3.9	1.9	0.54
2007	1335	6.1	3.6	1.9	0.54
2010	1306	5.3	3.1	1.6	0.54
Total	6977	5.9	3.7	1.9	0.56
Change (1997-2010)		-0.8**	-0.4*	-0.2*	-0.05***
% change		-12.53	-11.11	-11.25	

* p<0.10, ** p<0.05, *** p<0.01

Table 1.2: Mean landholding, total area planted and area allocated to plots with maize, by village population density (pooled sample – 1997-2010)

Quintile of village population density	N	Total landholding (acres)	Total area planted in main season (acres)	Maize	
				Area planted (acres)	Proportion of total area planted
1st (lowest)	1426	14.5	7.3	3.9	0.64
2	1394	5.67	3.6	1.7	0.53
3	1400	4.0	2.9	1.4	0.53
4	1418	3.2	2.5	1.2	0.52
5th (highest)	1339	2.2	1.9	1.0	0.56
Difference in mean (5th - 1st)		-12.3***	-5.4***	-2.9***	-0.07***

* p<0.10, ** p<0.05, *** p<0.01

As hypothesized, average landholding declined over the 13-year period by a higher percentage in the least than in the most densely populated villages (Table 1.3). This, however, was not the case for total area planted and area of plots with maize.

Table 1.3: Changes in mean landholding, total area planted and area allocated to plots with maize, by population density, (1997-2010)

Quintile of village population density	Total landholding (acres)	Total area planted in main season (acres)	Maize	
			Area planted (acres)	Proportion of total area planted
Absolute change (1997-2010)				
1st (lowest)	-2.5	-1.0	-0.5	-0.02
2	-1.0	-0.5	0.1	0.07
3	-0.1	0.0	-0.4	-0.16
4	0.0	-0.3	-0.1	-0.04
5th (highest)	-0.2	-0.2	-0.2	-0.08
% change				
1st (lowest)	-16.6	-13.6	-11.9	
2	-16.3	-14.2	9.6	
3	-3.2	0.8	-24.6	
4	-1.1	-10.9	-12.9	
5th (highest)	-8.2	-13.5	-17.3	

In Table 1.4, we categorize farm households into three groups according to landholding; less than 5 acres, 5-10 acres, and above 10 acres, and examine trends and changes in mean landholding, total area planted and area allocated to plots with maize. The average landholding in the <5 acres category increased by 12% overall, although between 2000 and 2010 average landholding consistently declined. Total area planted for this category also increased overall, but a consistent decline is observed between 2000 and 2010. Both average area of plots with maize and its share in total area planted declined over time. For the >10 acres category of landholding, total planted area and area of plots with maize increased marginally (by 7% and 9%, respectively). These results indicate that increasing maize production can only be through increased productivity even in areas where average landholding may be larger; there is little scope for area expansion for maize production.

Table 1.4: Levels of and changes in mean landholding, total area planted and area allocated to plots with maize, by landholding categories, 1997-2010

Landholding	Year	N	Total landholding (acres)	Total area planted in main season (acres)	Maize: Area planted (acres)	Maize: Proportion of total area planted
<5 acres	1997	968	2.0	2.0	1.1	0.60
	2000	931	2.4	2.4	1.3	0.56
	2004	868	2.3	2.1	1.0	0.54
	2007	878	2.3	1.9	1.0	0.54
	2010	915	2.2	1.8	0.9	0.53
	Pooled	4,560	2.2	2.0	1.1	0.56
	Change (1997-2010)		0.2	0.1	0.0	-0.05
	% change		12.4	3.4	-0.5	
5-10 acres	1997	330	6.8	4.0	2.0	0.56
	2000	319	6.8	4.5	2.4	0.55
	2004	359	7.0	5.0	2.5	0.52
	2007	286	6.9	4.4	2.3	0.55
	2010	263	6.7	4.0	2.1	0.57
	Pooled	1,557	6.8	4.4	2.3	0.55
	Change (1997-2010)		0.0	0.1	0.1	0.01
	% change		-0.6	1.9	4.8	
>10 acres	1997	201	24.4	10.4	5.0	0.55
	2000	191	22.3	11.6	5.9	0.59
	2004	169	25.2	10.7	5.0	0.55
	2007	171	24.0	11.2	5.9	0.57
	2010	128	24.1	11.1	5.4	0.55
	Pooled	860	24.0	11.0	5.4	0.57
	Change (1997-2010)		-0.3	0.7	0.4	0.01
	% change		-1.2	6.6	8.6	

A wide variation obtained in the land variables across agro-ecological zones (Table 1.5). Overall, average landholding was largest in the High potential maize (HPM) zone and smallest in the Western Highlands (WH) and Central Highlands (CH), the most densely populated of the zones. This pattern is also reflected in total area planted. It is worth noting that despite having the lowest yield levels of maize (shown later) among the agro-ecological zones, the Lowlands had among the highest average share of area of plots with maize in total area planted.

Over time, average landholding declined in all the zones except in WH, while average total area planted declined in most of the zones (Table 1.6). Average area of plots with maize also declined in all the zones except in the CL. Positive but small changes in the average share of plots with maize in total area planted were registered only in the HPM, CL and Western Transitional (WT) zones.

The patterns and trends in the land variables across agro-ecological zones provide an important insight to agricultural intensification. Because most rural households attempt to meet their own staple maize consumption needs through own production, maize production remains important to most smallholder farmers even in agro-ecological settings where maize productivity is relatively low. With declining farm sizes even in these agro-ecologies, it remains a question of whether efforts should be directed at increasing maize yield, or whether it would be more worthwhile to encourage higher value enterprises such as horticulture in such areas.

Table 1.5: Mean landholding, total area planted and area allocated to plots with maize, by zone (pooled sample – 1997 - 2010)

Zone	N	Total landholding (acres)	Total area planted in main season (acres)	Maize	
				Area planted (acres)	Proportion of total area planted
Coastal Lowlands	379	5.5	3.4	2.3	0.63
Eastern Lowlands	780	5.8	3.5	2.4	0.68
Western Lowlands	850	3.7	2.6	1.3	0.60
Western Transitional	792	5.6	4.1	1.5	0.42
High Potential Maize Zone	1,905	10.9	6.0	3.2	0.65
Western Highlands	742	2.3	1.9	0.9	0.53
Central Highlands	1,273	2.8	2.2	0.7	0.37
Marginal Rain Shadow	256	5.0	1.8	1.2	0.70
Overall	6,977	5.9	3.7	1.9	0.56

Table 1.6: Changes in mean landholding, total area planted and area allocated to plots with maize, by zone, 1997-2010

Zone	Total landholding (acres)	Total area planted in main season (acres)	Maize	
			Area planted (acres)	Proportion of total area planted
	Absolute change (1997-2010)			
Coastal Lowlands	-0.2	0.0	0.2	0.04
Eastern Lowlands	-2.0	-0.5	-0.6	-0.07
Western Lowlands	0.0	0.5	-0.1	-0.14
Western Transitional	-0.9	-1.1	-0.2	0.03
High Potential Maize Zone	-0.8	-0.5	0.0	0.01
Western Highlands	0.4	0.1	-0.1	-0.11
Central Highlands	-0.8	-0.6	-0.3	-0.07
Marginal Rain Shadow	-0.7	-0.2	-0.2	-0.05
Overall	-0.1	0.1	0.1	-0.03
	% change			
Coastal Lowlands	-4.2	1.4	9.6	
Eastern Lowlands	-31.0	-17.8	-27.3	
Western Lowlands	-1.0	22.9	-5.7	
Western Transitional	-15.6	-25.3	-16.4	
High Potential Maize Zone	-7.0	-8.8	-1.1	
Western Highlands	17.4	7.5	-13.8	
Central Highlands	-28.1	-24.7	-39.9	
Marginal Rain Shadow	-13.8	-10.8	-14.9	
Overall	-12.5	-11.1	-11.3	

It is a common practice among smallholder farmers in Kenya to plant maize together with other crops such as beans on the same plot. As shown in Table 1.7, planting maize alone on a plot was a rare practice overall; a maize plot was generally intercropped with at least one other crop. Besides, the average share of plots with intercropped maize in total maize area increased over time (by 6 percentage points). This would be expected since declining landholdings impose constraints on land available for production and encourages farmers to mix crops on the available cultivable land. While intercropping is generally a good agronomic practice, it has little benefits to maize yield if there are no beneficial interactions between maize and the other crop(s). Intercropping maize with legumes, such as is common in Kenya, is beneficial to maize as legumes are nitrogen fixers in the soil.

Table 1.7: Average share of monocropped and intercropped maize plots in total area planted and in total area of plots with maize, 1997 - 2010

Year	N	Monocropped maize area as % of …		Intercropped maize area as % of …		No. of other crops on intercropped maize plots
		Total area planted	Area under maize	Total area planted	Area under maize	
1997	1,499	7.54	14.85	51.70	85.15	1.6
2000	1,441	4.35	8.54	52.74	91.46	2.8
2004	1,396	4.40	9.05	50.03	90.95	2.4
2007	1,335	6.23	10.70	48.82	89.30	2.5
2010	1,306	4.30	8.12	50.30	91.88	2.2
Total	6,977	5.39	10.33	50.77	89.67	2.3
Change (1997 - 2010)		-3.23	-6.73	-1.40	6.73	0.7

1.4.2 Patterns and trends in land and labor productivity in maize production

In this section, we present and discuss patterns and trends in land productivity, family labor use rate and labor productivity (returns to family labor) in maize production. As explained earlier, the analysis is on the households' largest plots planted with maize (monocrop or intercrop) and for the years 2004, 2007 and 2010 for which labor data is available. We measure labor productivity in terms of net returns to family labor-day, adjusted for inflation. Net returns is computed as gross value of output for all crops planted with maize on the largest plot less the cost of variable inputs used on the plot, and adjusted for inflation using Fisher's ideal price index with 2004 quantities and district median prices as the base. We disaggregate the analysis by various household and spatial attributes to tease out relationships that might otherwise be masked by pooling the data.

Over time, average land productivity (net returns per acre) increased by 26% while family labor average use rate declined by 16% between 2004 and 2010 (Table 1.8). Average labor productivity (net returns to family labor) increased by 40%, a result of a combination of the increase in average land productivity and a decline in family labor average use rate. It is important to note that the decline in family labor average use rate was larger relative to the decline in the average planted plot size, indicating an increase in average land-labor ratio. The reduction in family labor use rate was compensated in part by increased use of hired labor, from an average of 17 to 19 person-days per acre, representing an increase of about 12% between 2004 and 2010.

Table 1.8: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) over time (mean)

Year	N	Planted plot size (acres)	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
2004	1153	1.5	10744	49.0	609
2007	1098	1.5	11484	39.1	933
2010	1005	1.3	13565	41.2	854
Total	3256	1.4	11864	43.2	794
% change (2004 - 2010)		-17.1	26.3	-16.0	40.3

Note: Net revenue values are in real terms

Across categories of landholding, average land productivity increased in the <5 acres and 5-10 acres categories but declined in the >10 acres category, while family labor average use rate declined for all the categories (Table 1.9). Average labor productivity also increased for all categories of landholding. The driving factor in the increase in average labor productivity for the 5-10 and >10 acres categories of landholding is the decline in family labor use rate rather than positive changes in land productivity.

Table 1.9: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by landholding (mean)

Landholding category	Year	N	Land productivity (Net revenue per acre (Ksh))	Family labor (Person-days per acre)	Labor productivity (Net revenue per family labor-day (Ksh))
<5 acres	2004	698	11489	58.8	379
	2007	726	12488	46.1	578
	2010	719	15016	46.5	601
	Total	2143	13011	50.4	521
	% change (2004 - 2010)		30.7	-20.8	58.7
5-10 acres	2004	307	9057	38.3	788
	2007	239	9957	29.2	1392
	2010	205	10302	31.9	1063
	Total	751	9683	33.6	1055
	% change (2004 - 2010)		13.8	-16.8	34.9
>10 acres	2004	148	10727	25.0	1325
	2007	133	8747	18.7	2047
	2010	81	8947	17.0	2574
	Total	362	9601	20.9	1870
	% change (2004 - 2010)		-16.6	-31.9	94.3

Note: Net revenue values are in real terms

Levels of and changes in average land and labor productivity and family labor use rate disaggregated by quintiles of population density are presented in Table 1.10. We observe several patterns. First, total average land productivity generally increased as population density increased. This relationship is much clearer in the non-parametric locally weighted scatterplot smoothing (lowess) regression analysis presented in Figure 1.3. It shows a concave relationship between land productivity and population density, indicating land productivity gains from agricultural intensification efforts by households as population increases and landholdings become smaller up to some point after which productivity gains diminish as population density increases (Muyanga & Jayne, 2014) . Over time, average land productivity increased for households in all the villages except those in the 1st quintile, mirroring the pattern observed in the disaggregation by landholding categories above.

Table 1.10: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by population density (mean)

Quintile of population density	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
1st (lowest)	2004	227	12449	32.0	1173
	2007	218	9419	23.8	1881
	2010	208	10475	27.3	1711
	Total	653	10809	27.8	1581
	% change (2004 - 2010)		-15.9	-15.0	45.8
2	2004	236	8947	43.9	519
	2007	219	11295	34.9	771
	2010	200	12442	40.9	733
	Total	655	10799	40.0	669
	% change (2004 - 2010)		39.1	-6.9	41.3
3	2004	226	9524	55.6	524
	2007	222	13119	42.9	814
	2010	198	13336	46.3	498
	Total	646	11928	48.4	615
	% change (2004 - 2010)		40.0	-16.7	-5.0
4	2004	231	10223	52.4	403
	2007	227	11551	42.9	770
	2010	206	15924	46.7	677
	Total	664	12446	47.4	614
	% change (2004 - 2010)		55.8	-11.0	67.9
5th (highest)	2004	233	12602	60.9	437
	2007	212	12017	50.9	425
	2010	193	15777	45.4	611
	Total	638	13368	52.9	486
	% change (2004 - 2010)		25.2	-25.5	39.7

Note: Net revenue values are in real terms

Secondly, as expected, family labor total average use rate increased with increase in population density. Over time, family labor average use rate declined in all the quintiles of population density, with the 2nd quintile registering the least and the 5th quintile the largest decline. The decline over time in family labor average use rate even in most densely populated villages suggests that households were devoting more labor to other activities, either/both farm or/and non-farm activities.

Lastly, there was an inverse relationship between total average labor productivity and population density, as also shown in the lowess regression analysis results in Figure1.4, indicating the need for faster growth in land productivity to sustain growth in returns to family labor as landholdings shrink due to increasing population.

Figure 1.3: Locally weighted scatterplot smoother (lowess) of land productivity of maize and village population density

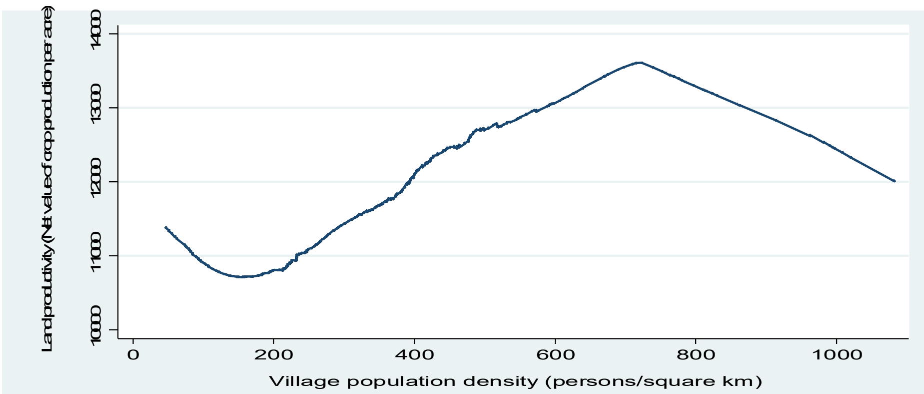
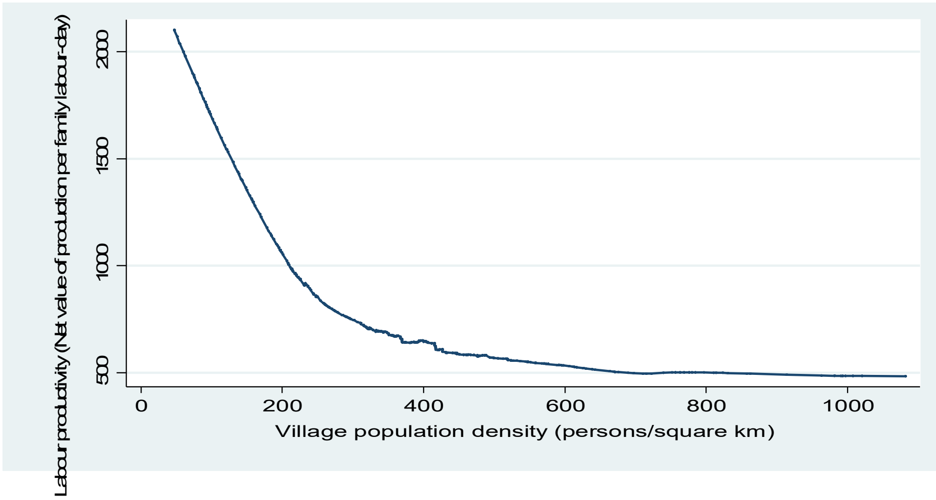


Figure 1.4: Locally weighted scatterplot smoother (lowess) of labor productivity of maize and village population density



A notable observation across gender of household head is that total average land productivity, family labor average use rate and labor productivity were higher in male- than in female-headed households (Table 1.11), indicating a generally higher productivity of maize in male-headed relative to that in female-headed households. Over time, both groups of households experienced positive changes in average land and labor productivity and a decline in family labor average use rate. It is important to note that 93% of female household heads did not have spouses and most (87%) were widows. This contrasts sharply with male household heads 75% of who had spouses. It implies that female-headed households in the sample lacked the support of one additional adult in their farming activities and may be the reason they generally had lower productivity compared to male-headed households.

The relationship between market access, represented by distance to paved road, and land and labor productivity and family labor use rate is presented in Table 1.12. On average, most remote households compared to their least remote counterparts had lower land and labor productivity and family labor use rate. Figure 1.5 and Figure 1.6 also show that land and labor productivity each generally has an inverse relationship with distance from a household to a tarmac road. This indicates that better transport infrastructure implies better access to both input and output markets, which may facilitate efficient use of inputs and family labor allocation, resulting in greater productivity of both land and family labor.

Table 1.11: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by gender of household head (mean)

Gender of household head	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (P*erson-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Female	2004	223	10085	45.5	507
	2007	263	10238	35.9	784
	2010	257	12131	39.3	732
	Total	743	10847	39.9	683
	% change (2004 - 2010)		20.3	-13.6	44.2
Male	2004	930	10902	49.8	633
	2007	835	11876	40.1	980
	2010	748	14058	41.8	896
	Total	2513	12165	44.2	827
	% change (2004 - 2010)		28.9	-16.1	41.5

Note: Net revenue values are in real terms

Table 1.12: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by distance to tarmac road (mean) (pooled sample – 2004 – 2007)

Quintile of distance to tarmac road	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (P*erson-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
1st (shortest)	828	12057	39.1	896
2	547	11821	43.7	804
3	689	12738	43.8	887
4	582	13073	44.0	894
5th (longest)	610	9502	47.1	446
Difference (5th-1st)		-2555	8.0	-450

Note: Net revenue values are in real terms

Figure 1.5: Locally weighted scatterplot smoother (lowess) of land productivity and distance to tarmac road

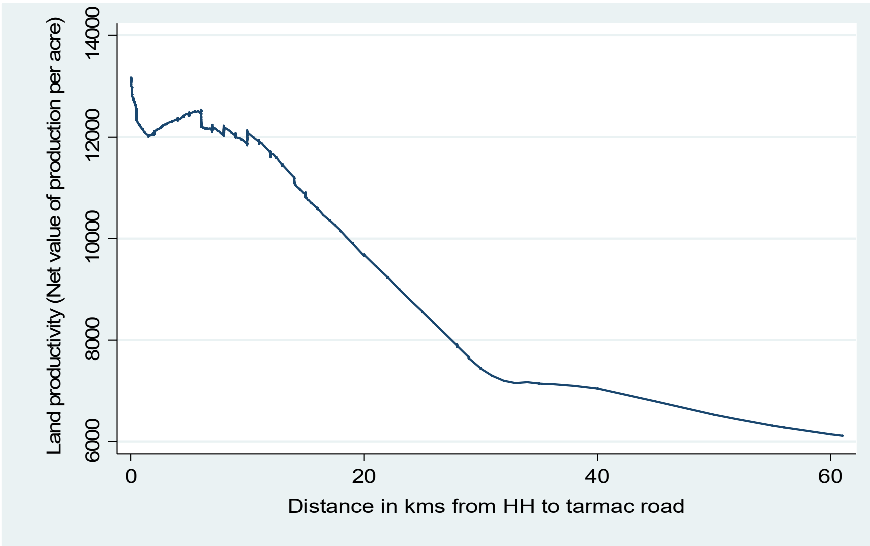
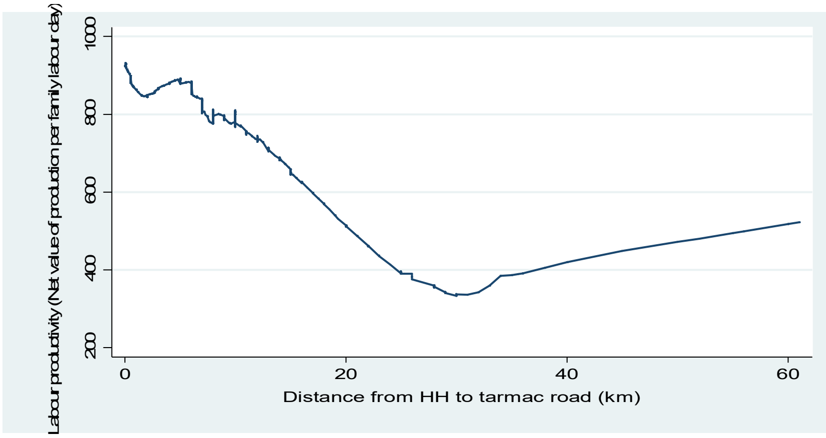


Figure 1.6: Locally weighted scatterplot smoother (lowess) of labor productivity and distance to tarmac road



With respect to cropping system, Table 1.13 shows that total average land productivity and family labor use rate were higher on intercropped than on monocropped plots, while total average labor productivity was higher on monocropped than on intercropped plots. Over time, average land and labor productivity increased while family labor use rate declined on intercropped plots. On monocropped plots, both average land productivity and family labor use rate declined, while labor productivity increased. It is worth noting that intercropped plots were significantly smaller than monocropped plots, on average. These patterns mirror the results observed previously that households with smaller landholdings had greater land productivity than those with larger landholding, and labor productivity was greater for households with larger landholding. Intensification of maize production involving intercropping with especially legumes may be useful in increasing land productivity.

Table 1.13: Land and labor productivity and family labor use rate on maize intercrop and monocrop plot (mean)

Cropping system of maize on plot	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Intercrop	2004	1059	10938	49.4	602
	2007	959	11994	39.8	893
	2010	892	14343	42.3	745
	Total	2910	12329	44.0	742
	% change (2004 - 2010)		31.1	-14.4	23.8
Monocrop	2004	94	8560	44.8	691
	2007	139	7967	34.0	1209
	2010	113	7430	32.5	1715
	Total	346	7953	36.5	1234
	% change (2004 - 2010)		-13.2	-27.5	148.2

Note: Net revenue values are in real terms

There were wide variations in average land and labor productivity and family labor use rate across agro-ecological zones (Table 1.14). The total average land productivity was highest in the CH, one of the most densely populated of the zones, and lowest in the CL. Family labor average use rate was in the WH and CH zones, which had the highest population density. Total average labor productivity was highest in the HPM zone, where family labor average use rate was also lowest, and lowest in CL.

Over time, average land productivity increased in all the zones except HPM, CL and MRS, while labor productivity declined only in CL. Family labor average use rate increased in all the zones except the EL.

These results have important implications. First, the reduction in average land productivity in the HPM zone, the most important maize producing region and where fertilizer and improved seed variety adoption on maize is nearly complete, implies a problem in low response of maize to input use. Research has shown that soil degradation is a major problem in Kenya's agriculture, and soil mining is a widespread concern not only in Kenya but in the entire Africa (Stoorvogel et al, 1993; The Montpellier Panel, 2013). Increasing maize production on degraded soils is not feasible without restoring soil fertility.

Secondly, both land and labor productivity were quite low in some zones, especially the Lowlands where, coincidentally, use of external inputs such as fertilizer and improved maize varieties is generally low. While it might be tempting to suggest that strategies to encourage widespread use of external inputs in these zones would be beneficial in raising maize productivity, it is more critical to recognize the necessity of agronomic practices that raise soil quality. Without building soil fertility through proper agronomic practices, encouraging use of external inputs such as fertilizers and improved seed varieties in these zones, despite their lower use rates, may not achieve the needed productivity growth.

Table 1.14: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by agro-regional zone (mean)

Agro-ecological zone	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Coastal Lowlands	2004	48	5603	59.7	194
	2007	57	8144	35.8	350
	2010	55	4724	40.5	151
	Total	160	6206	44.6	235
	% change (2004 - 2010)		-15.7	-32.2	-22.6
Eastern Lowlands	2004	129	6049	44.8	303
	2007	123	9950	28.8	684
	2010	127	13141	51.3	487
	Total	379	9692	41.8	489
	% change (2004 - 2010)		117.3	14.5	60.6
Western Lowlands	2004	117	4439	45.6	160
	2007	146	9280	33.6	793
	2010	140	7529	32.4	458
	Total	403	7267	36.7	493
	% change (2004 - 2010)		69.6	-29.0	186.5
Western Transitional	2004	149	9358	52.5	340
	2007	141	9839	54.4	352
	2010	143	13513	48.1	549
	Total	433	10887	51.7	413
	% change (2004 - 2010)		44.4	-8.4	61.5
High Potential Maize Zone	2004	361	13194	40.3	1134
	2007	309	10672	33.0	1552
	2010	215	11466	32.7	1654
	Total	885	11894	35.9	1406
	% change (2004 - 2010)		-13.1	-19.0	45.9

Table 1.14 (Cont'd)

Agro-ecological zone	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Western Highlands	2004	129	8992	58.9	282
	2007	136	11741	51.3	573
	2010	126	13249	41.8	729
	Total	391	11320	50.8	527
	% change (2004 - 2010)		47.3	-29.1	159.0
Central Highlands	2004	180	16817	58.5	664
	2007	151	19207	39.4	1158
	2010	173	25673	47.3	1025
	Total	504	20573	48.9	936
	% change (2004 - 2010)		52.7	-19.2	54.3
Marginal Rain Shadow	2004	40	11871	49.9	481
	2007	35	10976	46.2	653
	2010	26	5452	29.2	796
	Total	101	9909	43.3	621
	% change (2004 - 2010)		-54.1	-41.4	65.6

Note: Net revenue values are in real terms

Lastly, the high and increasing average land productivity in the WH and CH, the most densely populated zones, and WT and EL may suggest that sustainable intensification is taking place in the zones. It is important to note that average plot sizes in WH and CH zones are the smallest and a number of crops are intercropped with maize, which could also be among the reasons for the high average land productivity levels. All the same, strategies to ensure sustained momentum in increasing land productivity in these zones would be desirable, and lessons from these zones could inform the design of appropriate strategies for increasing land productivity in other less densely populated zones. It is also worth noting that there would be population density threshold beyond which no land productivity gains would be realized or there would be a decline in land productivity altogether (Josephson et al, 2014; Muyanga & Jayne, 2014; Ricker-Gilbert et al., 2014). Therefore, strategies to stimulate growth in non-farm sectors would certainly be necessary to ease population pressure on agricultural land.

Land and labor productivity (returns to land and to family labor) are also functions of use of other inputs and agricultural technologies. We would expect productivity of family labor to increase when farm operations are mechanized and when use of hired labor increases, if production increases thereof are high enough to offset mechanization and labor hiring costs. We explore these relationships in Tables 1.15 and 1.16. Total average land productivity and family labor use rate were higher for households that did not at all use hired labor than for those that used (Table 1.15). Average returns to family labor, on the other hand, was significantly higher for households that also used hired labor than for those that did not at all use, indicating that returns to family labor is much higher when a household is able to substitute hired for family labor. This implies that on average, households who can afford to hire labor for maize production may benefit from reallocating their labor from maize cultivation to activities with higher returns. Over time, both groups of households had increased average land productivity, with the group that did not use hired labor experiencing a larger increase. Average labor productivity increased more for households that did not use hired labor than for those that used hired labor.

In Table 1.16 are patterns and trends in land and labor productivity and family labor use rate by mechanized land preparation, a labor-saving technology. Data showed that mechanized land preparation was more common on larger plots, on average. While total average land productivity was significantly higher where land preparation was manual, total average labor productivity was higher where land preparation was mechanized. This indicates that mechanized land preparation does not really affect land productivity of maize but raises returns to family labor through reducing family labor requirement. Indeed, family labor average use rate was higher where mechanization was not used and much lower where it was used. Over time, average land and labor productivity increased for both groups of households. However, average land productivity increased only slightly for households that used mechanized land preparation. Family labor average use rate declined for both groups of households, but the percentage decline was larger for those that did not use mechanized land preparation.

1.15: Land and labor productivity and family labor use rate on maize plots (both monocrop and intercrop) by hired labor use (mean)

Use of hired labor	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Not used	2004	462	9709	67.6	200
	2007	446	12393	60.0	283
	2010	390	13835	56.8	296
	Total	1298	11871	61.7	257
	% change (2004 - 2010)		42.5	-16.0	47.9
Used	2004	691	11436	36.6	882
	2007	652	10862	24.8	1378
	2010	615	13394	31.3	1208
	Total	1958	11860	31.0	1150
	% change (2004 - 2010)		17.1	-14.5	36.9

Note: Net revenue values are in real terms

Table 1.16: Land and labor productivity and land-labor ratio on maize plots (both monocrop and intercrop) by mechanized land preparation (mean)

Use of mechanized land preparation	Year	N	Land productivity *(Net revenue per acre (Ksh))*	Family labor (*Person-days per acre*)	Labor productivity (*Net revenue per family labor-day (Ksh)*)
Not used	2004	584	11335	63.0	386
	2007	583	13693	48.1	656
	2010	565	16017	50.0	633
	Total	1732	13656	53.7	558
	% change (2004 - 2010)		41.3	-20.6	64.0
Used	2004	569	10137	34.6	838
	2007	515	8983	28.9	1247
	2010	440	10418	29.8	1138
	Total	1524	9828	31.3	1063
	% change (2004 - 2010)		2.8	-13.9	35.8

Note: Net revenue values are in real terms

We use expression (1.3) in the conceptual framework section to decompose percentage change in labor productivity into its components – percentage change in land productivity and in land-labor ratio – between two data points; 2004 and 2010 (Table 1.17). Overall, long-term labor productivity increased by 36% on average, resulting from a combination of an increase in land productivity of 21% and in land-labor ratio of 16%. The increase in land-labor ratio was because of a decline in family labor use rate. The fact that a little below half of the increase in returns to family labor is contributed by increase in land-labor ratio due to reduced family labor use rate attests to the need for faster growth in land productivity in Kenya's maize sector, and in agricultural sector in general.

Table 1.17: Mean percentage change in labor productivity of maize (both monocrop and intercrop) as the sum of percentage changes in land productivity and land-labor ratio, 2004-2010

		Land productivity	Land-labor ratio	Labor productivity
Overall		21.1	16.3	37.5
Farm size	<5 acres	31.8	21.3	53.1
	5-10 acres	4.5	19.0	23.5
	>10 acres	-42.2	65.8	23.6
Population density quintile	1st (lowest)	-34.5	26.7	-7.8
	2	34.9	2.8	37.7
	3	41.6	12.9	54.5
	4	45.6	4.8	50.5
	5th (highest)	19.1	31.2	50.3
Gender of household head	Female	26.0	19.8	45.8
	Male	20.7	15.2	35.9
Quintile of distance to tarmac road	1st (shortest)	11.4	12.4	23.8
	2	16.3	21.3	37.5
	3	29.0	15.5	44.5
	4	32.2	32.4	64.7
	5th (longest)	15.4	2.0	17.4
Cropping system	Intercrop	27.1	9.7	36.9
	Monocrop	-18.2	65.1	47.0
Agro-regional zone	Coastal Lowlands	-21.5	32.0	10.5
	Eastern Lowlands	83.6	-34.8	48.8
	Western Lowlands	59.4	38.3	97.7
	Western Transitional	41.1	4.5	45.6
	High Potential Maize Zone	-33.2	33.9	0.7
	Western Highlands	40.4	46.9	87.3
	Central Highlands	57.0	13.6	70.6
	Marginal Rain Shadow	-88.2	55.4	-32.8
Hired labor use	Not used	33.8	11.4	45.2
	Used	12.8	17.7	30.5
Mechanized land preparation	Not used	37.4	18.3	55.7
	Used	-0.3	22.0	21.8

The overall picture about the sources of change in labor productivity can mask important details regarding dynamics in the growth components across categories of farmers. For this reason, we disaggregate labor productivity change and its components by the various categories of households as in the analyses presented previously. By landholding, labor productivity increased for <5 acres category, contributed by positive change in land productivity (32%) and in land-labor ratio (21%). Again, it is important to note that the increase in land-labor ratio was because of a decline in family labor use rate. For the 5-10 acres category of landholding, labor productivity increased by 24%, contributed by 19% increase in land-labor ratio and only 5% increase in land productivity. Labor productivity also increased for the >10 acres category of landholding, because of an increase in land-labor ratio which outweighed the large decline in land productivity of 42%. These results indicate encouraging gains in production intensification by households with smaller landholdings, but a cause for concern about the lack of yield gains on larger farms, which are often relied upon for producing marketable surpluses. Therefore, it should not be a surprise that there is an increasing deficit in domestic supply of maize in Kenya (Kirimi et al., 2011).

The patterns in labor productivity changes observed across landholding categories are reflected in the quintiles of population density. Households in the least densely populated villages had a decline of 8% in labor productivity, contributed by a decline in land productivity of 35% and positive change of 27% in land-labor ratio. On average, labor productivity increased for households in the second to fourth quintiles of village population density, more because of positive change in land productivity and less because of increase in land-labor ratio. For the fifth quintile of population density, the increase in labor productivity was largely contributed by positive change in land-labor ratio. It is important to keep in mind that the positive change we observe in land-labor ratio was because of a decline in family labor use rate on maize plots and not increase in area planted with maize, which declined over time.

Disaggregation by gender of the head of household shows positive change in labor productivity for both male and female-headed households, because of increase in both land productivity and land-labor.

Across quintiles of distance to paved road, an indicator of market access, labor productivity increased for all the households but the increase was smallest for households furthest from paved roads. This was due to a 15% increase in land productivity and 2% increase in land-labor ratio.

While intercropped plots registered positive change in labor productivity due to increases in land productivity and land-labor ratio, monocropped maize plots had labor productivity increase because of a large increase in land-labor ratio that outweighed the negative change in land productivity.

A remarkable result across agro-ecological zones is the decline or small increase in labor productivity in the HPM, CL and MRS zones, because of a decline in land productivity. Being the most important maize producing region of Kenya, the dwindling maize yield in HPM zone indicates the urgent need for ways to address the problems associated with low land productivity.

Both households that used and those that did not use hired labor on maize production had positive change in labor productivity, because of positive change in land productivity and in land-labor ratio. Disaggregation by type of land preparation shows that households that used mechanized land preparation had a smaller increase in labor productivity because of a decline in land productivity. It is worth noting that mechanized land preparation, larger landholdings and low population density are highly correlated hence similar patterns in labor and land productivity changes.

1.5 Conclusions and implications for policy

Kenya's agricultural productivity has generally been low and has stagnated over time. Aggregate yield of maize, the most important staple crop, although showing a gradual upward trend in the recent past, has a general downward trend over the past two and a half decades. In addition, farm sizes have been shrinking over time as rural population densities rise and land sub-division occurs. These have made Kenya increasingly dependent on food imports for its national food security. The country is also foregoing opportunities to generate strong farm-nonfarm growth multipliers and shifts in the composition of the labor force for economic transformation. This study focused on maize and aimed to determine whether Kenya's agricultural sector is changing in ways that are promoting or retarding farm labor productivity and to understand the association between farm labor productivity and population density, land scarcity and market access. We measured labor productivity in terms of returns to family labor, which we consider most appropriate for profitability of farming for agricultural households.

Six key results have emerged. First, average landholding declined by over 12% in 13 years while average area planted declined by 11%. Maize is the single most important crop, occupying over 50% of total area planted. Over time, the average area of plots with maize for a household declined by 11%.

Second, long-term labor productivity (returns to family labor) increased by 38% overall, contributed by positive changes in land-labor ratio of 16% and land productivity of 21%. The increase in land-labor ratio was because of a decrease in family labor use rate on maize plots. The decline in family labor use rate was accompanied by increased use of hired labor. While many factors that the scope of the analysis in this paper is incapable of identifying could explain the decline in family labor use rate, one plausible explanation is that farm and non-farm activities of higher returns might be attracting labor of smallholder farm families, making them to reduce their labor in lower-value maize production.

Third, there was an inverse relationship between landholding and land productivity. The inverse relationship was also observed between landholding and family labor use rate. Over time, labor productivity increased substantially for the <5 acres category of landholding, contributed by increase in land productivity (32%) and in land-labor ratio (21%). The labor productivity increase for the 5-10 and >10 acres categories of landholding was mainly because of increase in land-labor ratio rather than in land productivity. In fact, land productivity declined by 42% for the >10 acres category of landholding.

Fourth, average land productivity and family labor use rate had direct relationships with population density while labor productivity was inversely related with population density. Households in the 1st quintile of population density (least densely populated villages) had a decline of 8% in labor productivity, contributed by a decline of 35% in land productivity and an increase of 27% in land-labor ratio. Labor productivity increased for households in all the other quintiles, more because of increase in land productivity than the increase in land-labor ratio, except for the 5th quintile (most densely populated) where the contribution of land productivity increase was lower than that of land-labor ratio increase.

Fifth, most remote households, in terms of distance to a paved road, compared to their least remote counterparts had lower land productivity, higher family labor use rate and lower labor productivity, on average. Although labor productivity increased for both households over time, the increase was smallest for the households furthest from paved roads, because of smaller increases in land productivity and land-labor ratio.

Lastly, there was a remarkably small increase in labor productivity (0.7%) in the HPM zone, Kenya's most important maize producing region, because of a 33% decline in land productivity.

These results provide important insights about Kenya's farming sector in general and smallholder agriculture in particular. First, the shrinking landholdings and the subsequent decline in area planted among smallholder farmers puts the onus on strategies to achieve agricultural growth through agricultural intensification and productivity growth. The sheer amount of area under maize means that the crop remains fundamental to Kenyan agriculture and economy and land productivity of maize will need to go up substantially for overall agricultural labor productivity to rise appreciably over time. This means that we cannot just focus on diversifying into other crops – that may be necessary but certainly not sufficient. Kenyan policy makers must also figure out how to raise maize yield to raise the returns to labor in Kenyan agriculture overall. One of the major causes of low and stagnated agricultural yields in Kenya is soil infertility (Government of Kenya, 2014; Tittonell et al., 2008; Marenya & Barrett, 2007). Therefore, raising maize yield will require addressing the soil infertility problem.

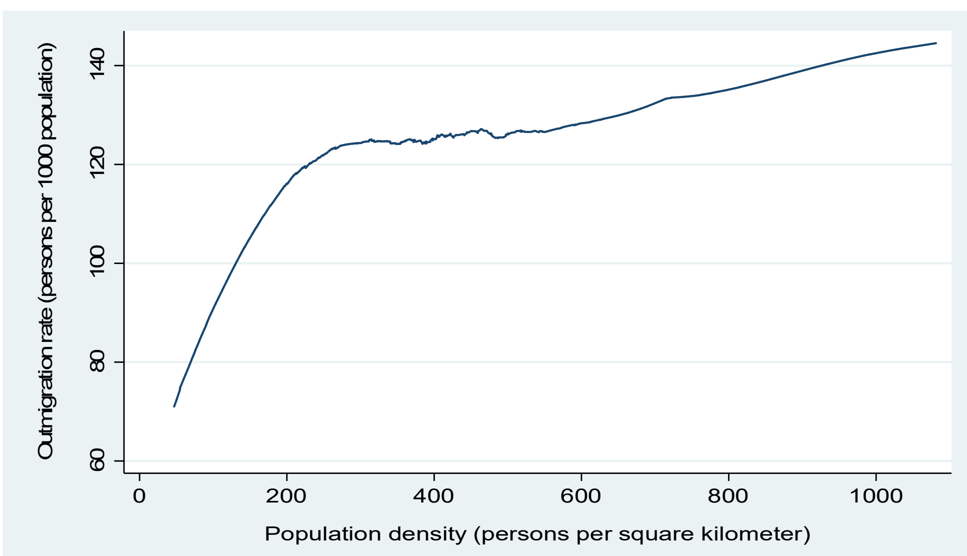
Secondly, we have shown that labor productivity increased in areas with smaller landholdings and higher population density because of increase in land productivity, and declined or only slightly increased in areas with larger landholdings and lower population density because of a decline in land productivity. This suggests encouraging gains in production intensification by households with smaller landholdings, but a cause for concern about the lack of productivity gains on larger farms, which are often relied upon for producing marketable surpluses. Therefore, addressing causes of declining yield on larger farms should be a priority.

Lastly, land-labor ratio increased even in villages with high population densities because households reduced family labor use rate on maize. On one hand this suggests that households may be devoting more labor to other activities, either/both farm or/and non-farm, that potentially provide much higher returns to their labor than cultivating maize. Shifting of the labor force involving movement of labor from farm to non-farm activities over time and accompanied by increased land productivity can be viewed as part of structural transformation process, which will by construct raise labor productivity in agriculture. This process appears to be taking place in some high densely populated villages where both land-labor ratio and maize yield increased. On the other hand, reduction of family labor in maize production might be based on "push" factors. Haggblade et al (2005) note that labor can move from agriculture to rural non-farm sector characterized by low-return activities because of low labor productivity in agriculture, low opportunity cost of labor and declining real income of households. In the light of this argument, reduction in family labor use rate on maize production may also be because of low or declining maize yield. In such scenario, it is essential that underlying causes of low land productivity be addressed to raise labor productivity.

APPENDIX

Figure A1.1 shows the relationship between village outmigration rate of the population 15 – 40 years of age. Outmigration rate is computed as the ratio of the number of persons aged 15- 40 years that migrated out of the village to the number of village residents in that age bracket, multiplied by 1000. The rate is thus interpreted in terms of number of out-migrants per 1000 population in the 15 – 40 years age bracket in a village. The figure shows a general positive association between village outmigration rate and population density. Marginal effects from a simple OLS regression of village outmigration rate on population density and its quadratic term shows that on average, and without accounting for other factors that might affect outmigration rate, additional population density of 100 persons per square kilometer in a village is associated with an increase in migration rate of 5 persons per 1000 population in the 15 - 40 years age bracket.

Figure A1.1: Relationship between youth outmigration rate and village population density (lowess regression using pooled sample – 2000-2010)



REFERENCES

REFERENCES

Argwings-Kodhek, G., Jayne, T. S., Nyambane, G., Awuor, T., & Yamano, T. (1999). *How can micro-level household information make a difference for agricultural policy making? Selected examples from the KAMPAP survey of smallholder agriculture and non-farm activities for selected districts in Kenya*. Nairobi. Retrieved from http://tegemeo.org/index.php/resources/publications/168-technical-reports/521-how-can-micro-level-household-information-make-a-difference-for-agricultural-policy-making-selected-examples-from-the-kampap-survey-of-smallholder-agriculture-and-non-farm-activ

Auer, L. (1966). Labour Productivity in Agriculture—a Conceptual Analysis. *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie*, *14*(1), 29–39. https://doi.org/10.1111/j.1744-7976.1966.tb02373.x

Barbier, E. B., & Hochard, J. P. (2014). *Land degradation, less favoured lands and the rural poor: A spatial and economic analysis*.

Boserup, E. (1965). The conditions of agricultural growth. *Population Studies*, *20*(1), 1–108. https://doi.org/10.2307/2172620

Chamberlin, J. (2013). *Market access and smallholder development in Kenya and Zambia*. Michigan State University.

FAOSTAT. (2014). FAOSTAT. *Food and Agricultural Organization of the United Nations*. Retrieved from http://data.fao.org/ref/262b79ca-279c-4517-93de-ee3b7c7cb553.html?version=1.0

Government of Kenya. (2014). *Soil suitability evaluation for maize production in Kenya soil suitability evaluation for*. Nairobi, Kenya.

Haggblade, S., & Hazell, P. B. R. (Eds.). (2010). *Successes in African Agriculture*. International Food Policy Research Institute.

Haggblade, S., Hazell, P., & Reardon, T. (2005). The rural nonfarm economy: Pathway out of poverty or pathway in? In *International Food Policy Research Institute. The future of small farms. Proceedings of a research workshop*. Wye, UK.

Horneck, D. A., Sullivan, D. M., Owen, J. S., & Hart, J. M. (2011). *Soil Test Interpretation Guide* (EC No. 1478). Corvallis, OR.

Jayne, T. S., & Rashid, S. (2013). Input subsidy programs in sub-Saharan Africa: A synthesis of recent evidence. *Agricultural Economics*, *44*, 547–562. https://doi.org/10.1111/agec.12073

Johnston, B. F., & Mellor, J. W. (1961). The role of agriculture in economic development. *The American Economic Review*, *51*(4), 566–593.

Josephson, A. L., Ricker-Gilbert, J., & Florax, R. J. G. M. (2014). How does population density influence agricultural intensification and productivity? Evidence from Ethiopia. *Food Policy*, *48*, 142–152. https://doi.org/10.1016/j.foodpol.2014.03.004

Kirimi, L., Sitko, N., Jayne, T. S., Karin, F., Muyanga, M., Sheahan, M., … Bor, G. (2011). *A farm gate-to-consumer value chain analysis of Kenya's maize marketing system* (Working Paper No. 44). Nairobi.

Lipton, M. (2005). *The family farm in a globalizing world: The role of crop science in alleviating poverty* (2020 Discussion Paper No. 40). Washington, DC.

Marenya, P. P., & Barrett, C. B. (2007). Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. *Food Policy*, *32*(4), 515–536.

Muyanga, M., & Jayne, T. S. (2014). Effects of rising rural population density on smallholder agriculture in Kenya. *Food Policy*, *48*, 98–113. https://doi.org/10.1016/j.foodpol.2014.03.001

Olwande, J., Smale, M., Mathenge, M. K., Place, F., & Mithöfer, D. (2015). Agricultural marketing by smallholders in Kenya: A comparison of maize, kale and dairy. *Food Policy*, *52*, 22–32. https://doi.org/10.1016/j.foodpol.2015.02.002

Otsuka, K., & Larson, D. F. (Eds.). (2013). *An African Green Revolution: Finding ways to boost productivity on small farms*. Dordrecht Heidelberg New York London: Springer. https://doi.org/10.1007/978-94-007-5760-8

Pingali, P., Bigot, Y., & Binswanger, H. P. (1987). *Agricultural mechanization and the evolution of farming systems in sub-Saharan Africa*. *Economic Geography* (Vol. 64). The Johns Hopkins University Press. https://doi.org/10.2307/144242

Ricker-Gilbert, J., Jumbe, C., & Chamberlin, J. (2014). How does population density influence agricultural intensification and productivity? Evidence from Malawi. *Food Policy*, *48*, 114–128. Retrieved from http://dx.doi.org/10.1016/j.foodpol.2014.02.006

Stoorvogel, J. J., Smaling, E. M. A., & Janssen, B. H. (1993). Calculating soil nutrient balances in Africa at difference scales. *Fertilizer Research*, *35*, 237–235. https://doi.org/10.1007/BF00750642

The Montpellier Panel. (2013). *Sustainable intensification: A new paradigm for African agriculture*. London.

The World Bank. (2008). *World Development Report 2008: Agriculture for Development*. Washington, DC.

Tittonell, P., Corbeels, M., Wijk, M. T. Van, Vanlauwe, B., & Giller, K. E. (2008). Combining Organic and Mineral Fertilizers for Integrated Soil Fertility Management in Smallholder Farming Systems of Kenya: Explorations Using the Crop-Soil Model FIELD. *Agronomy*

Journal, *100*(5), 1511–1526.

Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, *143*, 76–90. https://doi.org/10.1016/j.fcr.2012.10.007

Wambugu, S., Kirimi, L., & Opiyo, J. (2011). *Productivity trends and performance of dairy farming in Kenya* (Working Paper No. 43). Retrieved from http://tegemeo.org/index.php/resources/publications/166-working-papers/434-productivity-trends-and-performance-of-dairy-farming-in-kenya.html

Yeboah, K. F., & Jayne, T. S. (2016). *Africa's evolving employment structure* (MSU International Development Working Paper No. 48).

CHAPTER 2: FARMERS' PERCEPTIONS OF SOIL FERTILITY AND ADOPTION OF SOIL FERTILITY IMPROVEMENT PRACTICES

2.1 Introduction and Study Objectives

Soil infertility has been identified as a fundamental cause of low agricultural productivity in sub-Saharan Africa (Stoorvogel & Smaling, 1998; Sanchez et al., 1997), with nutrient depletion cited as the major cause for the widespread soil infertility (The Montepellier Panel, 2013; Smaling et al., 1997; Stoorvogel & Windmeijer, 1993). Stoorvogel & Windmeijer (1993) point out that cropping intensity and land management are significant determinants of the rate of nutrient depletion in soils once land use changes from natural vegetation to agriculture.

It is widely acknowledged that continuous cultivation of land without soil amendments in form of nutrient replenishment and organic matter addition degrades soils and lowers crops response to external inputs such as fertilizers (Sanchez et al., 1997; Lal, 2006; Tittonell & Giller, 2013). Extractive land management practices deplete soil organic carbon (Lal, 2006), the main component of soil organic matter. Soil organic carbon is essential for a range of soil chemical, physical, and biological properties and is an overall indicator of soil health (Okalebo, Gathua, & Woomer, 2002;Lal, 2006; Horneck et al., 2011). Not least, soil organic carbon influences soil moisture and nutrient holding capacities, microbiological activity, and soil structure and other physical properties (Okalebo, Gathua, & Woomer, 2002; Lal, 2006; Horneck et al., 2011). Soil pH, an important chemical property that influences soil nutrient availability to plants (McCauley, Jones, & Olson-Rutz, 2017; IIASA/FAO, 2012) and is a measure of soil acidity, is also influenced by land management practices. Although it is a naturally occurring process, Kunhikrishnan et al

(2016) point out that soil acidification on farms is mainly due to inappropriate use of chemical fertilizers, in addition to improper management of organic materials on the farm. Continued application of acidifying fertilizers such as DAP to soils with low pH and low organic matter can further degrade the soil. When soils are degraded, raising productivity can be possible only through management practices that add organic matter and nutrients and ameliorate soil acidity to restore fertility (Lal, 2006; Chivenge et al., 2011; Kunhikrishnan et al., 2016).

In Kenya, studies have shown a remarkable increase in farmer adoption of fertilizer and improved maize seed over time (e.g. Ariga & Jayne, 2009; Smale & Olwande, 2014), yet no growth in maize yield has been visible over the last quarter century. Soil degradation because of continuous cultivation without sufficient replenishment of soil nutrients, organic matter or both is one of the main reasons for low agricultural productivity in general and poor maize yield in particular (Government of Kenya, 2014; Tittonell et al., 2008; Marenya & Barrett, 2007). This implies that continued use of fertilizers and improved maize varieties on their own cannot achieve the needed growth in yield. It requires change of the current disproportionate emphasis by the government, development agencies and non-governmental organizations on promoting chemical fertilizer use to increase maize yield. There is a need to correct for the deficiency in soil fertility through encouraging use of sustainable soil management practices.

Lal (2006) emphasizes the importance of sustainable management of soil resources to enhancing soil fertility. Sustainable management includes agronomic practices that protect soil from erosion and conserve moisture (e.g. zero and minimum tillage, mulching, terracing, contour farming, use of soil bunds), and that recycle organic matter and add nutrients into the soil (e.g. use of manure, compost and crops residues, and judicious application of chemical fertilizers). These practices, apart from relying only on chemical fertilizers, enhance soil fertility by adding organic matter to the soil, which increases soil organic carbon and, depending on some other factors, buffers soil pH change (McCauley et al., 2017). They also improve water and nutrient use efficiency. Therefore, encouraging widespread and sustained use of sustainable management practices that enhance soil fertility should be a priority in Kenya if agricultural productivity growth is to be realized.

Many studies have explored factors that influence farmers' adoption and use of agricultural technologies and agronomic practices in Kenya. For example, Wainaina et al. (2016) used a multivariate probit model to study factors that affect adoption of high yielding varieties of maize seed, inorganic fertilizers and a range of soil management technologies (zero tillage, use of crop residues, manure use and terracing and soil bunds). Using a similar analytical approach, Kamau et al. (2014) studied adoption of soil conservation practices, inorganic fertilizers and a combination of other soil fertility management practices (mulching, manure use, use of compost and planting legumes). Marenya & Barrett (2009) applied a switching regression model to study factors that affect use of inorganic fertilizer. Switching regression was used to determine how the factors affect fertilizer use below and above a certain threshold level of soil carbon content. In another study, Marenya & Barrett (2007) analyzed factors that affect adoption and dis-adoption of inorganic fertilizer, manure, maize crop residue and agroforestry. Odendo et al. (2009) applied a logit model to analyze adoption of inorganic fertilizer, manure, compost, and a combination of all three. These studies have a common feature that they applied econometric methods to household- and/or plot-level data to identify variables that explain adoption of agricultural technologies and agronomic practices. They emphasize quantitative factors that affect adoption to inform policy prescriptions for farmers' adoption of agricultural technologies and agronomic practices. The often-considered factors include social, demographic and economic characteristics of the farmer, institutional and market conditions and environmental factors.

We can expect that farmers' decisions regarding adoption of technologies and agronomic practices that improve soil fertility may be influenced by, among other factors, what they think or know about the fertility condition of their soil, their perceptions and knowledge about the technologies and practices, and their ability to access the technologies and apply the practices. Few quantitative studies exist about the correspondence of farmers' perceptions about the fertility conditions of their soils and measured soil fertility, and how those perceptions influence adoption (or use) of soil fertility management technologies and agronomic practices. Two studies that have conducted such analysis are Marenya et al (2008) and Berazneva et al. (2016). In an econometric framework using data on a sample of plots in a region in Western Kenya, Marenya et al (2008) analysed the relationship between farmers' perceptions about soil fertility and laboratory measure of soil carbon, which they used as a measure of soil fertility. The study also analyzed the relationship between farmers' subjective perceptions about impacts of fertilizer and statistically derived marginal product of nitrogen use on maize and beans. The study concluded that observed crop yields influenced farmers' perceptions about soil fertility and impacts of fertilizer. The study by Berazneva et al. (2016) used data from two locations in western Kenya and a nationally representative sample in Tanzania to identify correlates of farmers' reported perceptions about soil quality and compare the perceptions with measured soil fertility. Similar to Marenya et al (2008), Berazneva et al. (2016) found that farmers' perceptions about soil quality was driven by crop yields. Despite the dearth of studies in this area, it remains germane to understand farmers' perceptions about soil fertility conditions and their influence on use of soil management practices to support policy and extension efforts to address the acute problem of soil infertility and attendant low agricultural productivity in Kenya.

This study has two objectives. The first is to assess maize farmers' perceptions about the fertility of their soils and compare these with measured soil fertility based on results from scientific test of soil chemical properties. The second is to investigate farmers' adoption (use) of soil fertility management practices, with a focus on the influence of their perception of soil fertility. The research questions guiding the study are as follows:

(a) What are farmers' perceptions about the fertility of their soils? How do farmers' perceived and measured soil fertility compare? What informs farmers' subjective judgement of soil fertility?

(b) What is the relationship between farmers' perceptions about soil fertility and their adoption of soil fertility management practices?

This study differs in several ways from Marenya et al (2008), Berazneva et al. (2016) and the studies highlighted above that analyzed factors that affect farmers' adoption and use of agricultural technologies and agronomic practices in Kenya. First, in the adoption analysis, we account for cognitive aspects underlying farmers' decision to apply various soil fertility management practices by including their perception about the fertility condition of their soils among the explanatory variables. Secondly, different from Marenya et al (2008), we use a soil fertility measure that combines soil organic carbon, nitrogen and pH, based on recommended thresholds for optimal maize growth in Kenya. While Berazneva et al. (2016) also used the three soil chemical properties to construct measured soil fertility index, the threshold values they used to construct the index were not based on recommended thresholds for optimal maize growth in Kenya despite the study's focus on maize plots. Thirdly, different from Marenya et al (2008), we analyze not only the relationship between farmers' perceptions and measured soil fertility, but also the relationship between farmers' perceptions and individual measured chemical and physical attributes of the soil. This is important for identifying soil properties that might be underlying farmers' perceived soil fertility status. Lastly, our sample covers the primary maize growing areas of Kenya, making the study most relevant to search for potential solutions to the problem of low and stagnated maize productivity because of soil infertility.

2.2 Methods and Data

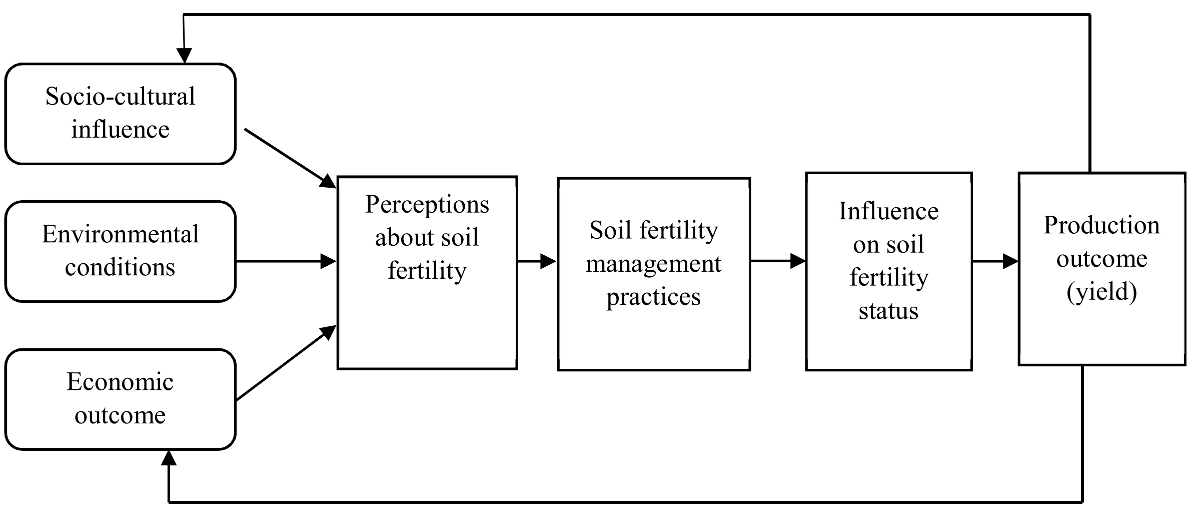
2.2.1 Conceptual Framework

Econometric methods have dominated studies about farmers' use of soil fertility management technologies and practices to generate explanations regarding factors that hinder or promote their use. The adoption studies highlighted in the previous section for Kenya and many others have often used farmer attributes such as gender and education, plot and farm level characteristics, and institutional (market and policy) and environmental contexts as covariates in regression analyses to arrive at the explanations. While the analyses are certainly useful in understanding farmer adoption behavior, a limitation in many of them is lack of information about cognitive aspects that underlie farmers' decision-making regarding use of technologies and practices. Yet, it is hardly disputable that farmers respond to a phenomenon, in this case soil fertility condition, based on, among other factors, what they know or believe about it.

Although not common in studies about what influences farmers' agricultural decisions, there is increasing recognition and evidence that farmers' values, perceptions and beliefs influence decisions concerning soil conservation and fertility management (e.g. Pender & Kerr, 1998; Enyong et al., 1999; Eckert & Bell, 2005; Asafu-Adjaye, 2008; Vignola et al., 2010; and Turner et al., 2014). The research in this paper considers farmers' perception about soil fertility and its influence on decisions to apply selected fertility management practices. It uses two separate analyses, but in a way that generates a unified understanding of the importance of farmers' perception to soil fertility management in maize production in Kenya.

We suggest that a range of factors may influence farmers' perceptions about the fertility of their soils (Figure 2.1). Such factors may include socio-cultural influence (e.g. accumulated experience in farming), environmental conditions (e.g. soil properties) and economic outcomes (e.g. crop yield) (Halbrendt et al., 2014; Marenya et al., 2008; Desbiez et al., 2004; Corbeels et al., 2000). These perceptions inform decisions regarding soil management practices to apply to influence fertility for a desired production outcome (yield). The resulting outcome may then feedback to inform future perceptions through economic outcomes and socio-cultural influence.

Figure 2.1: Conceptual framework for understanding farmers' perceptions about soil fertility and its influence on fertility management practices (Adapted from Halbrendt et al. (2014))



The first analysis entails estimating the extent of agreement between farmers' perception about (subjective judgement of) the fertility of their soils and fertility as measured from scientific test results of soil chemical properties. The main idea is that farmers' subjective conclusion about the fertility condition of their soils might significantly differ from measured soil fertility. When this happens, it would have undesirable implications about farmer knowledge of the fertility needs of their soils and the appropriateness of soil management practices that they prioritize.

The second analysis examines the relationship between farmers' perceived soil fertility and their adoption (use) of soil fertility management practices. It applies an econometric modelling framework for agricultural technology adoption.

Choice of soil fertility management practices is largely an economic decision. We motivate the analysis using the agricultural household model originally developed by Singh et al. (1986) and later extended by de Janvry et al. (1991) to explain farm household's decision-making in settings with semi-subsistence production and market failures. We assume non-separability of household production and consumption decisions because agricultural households in Kenya are generally semi-subsistent and often face numerous market imperfections. Input markets, for example for inorganic fertilizers, are imperfect while markets for organic materials such as manure, compost and crops residues are generally missing. Because of this, prices that guide farmers' choices of management practices are endogenously determined and depend on both household-specific characteristics and characteristics of markets. Therefore, farmers' objective is assumed to be utility maximization rather than profit maximization.

We conceptualize adoption of fertility management practices in the context of random utility framework as explained in Ben-Akiva and Lerman (1985) and applied in Marenya & Barrett (2007), Ali & Abdulai (2010) and Kassie et al. (2011). Using the explanation in Ben-Akiva and Lerman (1985) as the basis, we argue that when a farmer is faced with a feasible set of discrete alternative management practices to choose from, the farmer will select that which generates the greatest perceived utility. Assuming a linear utility function for practice g for household i, utility is represented as a function of observed and measured characteristics specific to a household and plot, institutional factors, and environmental conditions ($\boldsymbol{V}_{ig}$), and a stochastic term (ε_{in}). The stochastic term captures the difference between the estimated and actual utility and accounts for unobserved and unmeasured factors not in $\boldsymbol{V}_{ig}$ that affect utility. The probability of household i adopting practice g among a set of G_i alternative practices available to it is equal to the probability that the utility of practice g, U_{ig}, is at least as great as the utility of all other alternative practices in G_i. That is,

$$P_{(ig)} = P\left(U_{ig} \geq U_{jg}, \forall j \in G_i\right) = P\left(\boldsymbol{V}_{ig} + \varepsilon_{ig} \geq \boldsymbol{V}_{jg} + \varepsilon_{jg},\ \forall j \in G_i\right) \qquad (2.1).$$

The farmer faces a discrete choice of a soil fertility management practice that yields the greatest utility, but we cannot observe the utility. What we do observe is whether a farmer has implemented a practice. We thus proceed by specifying a latent variable function for practice g for household i as:

$$T_{ig}^* = \boldsymbol{V_{ig}\beta} + \varepsilon_{ig}\ , \qquad T_{ig} = 1[T_{ig}^* > 0] \qquad (2.2),$$

where the latent variable T_{ig}^* represents net benefit to household i from adopting practice g, T_{ig} is an indicator variable that equals 1 if household i adopts practice g and 0 if not, and $\boldsymbol{V}$ is a vector of household, farmer and plot characteristics, institutional factors and environmental conditions that explain heterogeneity in adoption of the practice. The vector $\boldsymbol{\beta}$ represents parameters to be estimated and ε_{ig} is idiosyncratic error term. Among the farmer characteristics is our variable of interest: farmer's perception about the fertility of soil of the plot in question.

Literature on adoption of the ‘Green Revolution’ technologies, for example inorganic fertilizers, high yielding seed varieties and pesticides, have comprehensively documented factors that affect adoption, including risk and uncertainty, capital constraints, credit constraints, farm size, constraints in supply of complementary inputs and land tenure arrangement (Feder et al., 1985; Kelly et al., 2003; Foster & Rosenzweig, 2010; Udry, 2010; Jack, 2011) . A distinctive feature of this literature is that it does not explicitly include farmers’ perceptions among the variables deemed to explain technology adoption decisions.

Another strand of literature has focused on adoption of soil management practices. The practices include those that conserve water and soil and those that improve soil fertility, which may include use of green revolution technologies, specifically mineral fertilizers. For example, Knowler & Bradshaw (2007) review studies on adoption of conservation agriculture, which is a package of soil and water conservation practices. They group the explanatory variables used in the studies into characteristics of the farmer, household and farm, and exogenous factors (e.g. market prices, extension services, and social capital). Farmer characteristics include perceived risk and attitudes towards conservation. Studies by Kamau et al. (2014), Marenya & Barrett (2009) and Odendo et al. (2009) examine factors that affect adoption of a range of soil management technologies, such as zero tillage, use of crop residues, mulching, manure use, use of compost, use of inorganic fertilizers, planting legumes, terracing, soil bunds and agroforestry. These studies use in the adoption equations a range of explanatory variables, including plot biophysical characteristics, farm household socio-economic characteristics, institutional factors such as proximity to markets, and environmental conditions. However, the studies as well as those reviewed by Knowler & Bradshaw (2007) do not take into account the role of farmers’ perceptions about the condition of their soils, yet these would be expected to influence their decisions about adoption of soil management practices.

As explained at the beginning of this section, there is increasing recognition that farmers' perceptions matter in their decisions concerning soil fertility management and conservation (e.g. Pender & Kerr, 1998; Enyong et al., 999; Eckert & Bell, 2005; Asafu-Adjaye, 2008; Vignola et al., 2010; Turner et al., 2014). Failure to incorporate farmers' perceptions in adoption decisions might provide an incomplete account of important factors affecting those decisions. This is especially important for Kenya, where there is urgent need for soil fertility improvement to raise the generally low and stagnated or declining agricultural productivity.

2.2.2 Analytical Strategy

2.2.2.1 Farmers' perceived and measured soil fertility

We use three approaches to compare farmers' perceived and measured soil fertility. First, we apply chi-square test for independence to determine whether there is a relationship between farmers' perceived and measured soil fertility. Secondly, we use interrater agreement technique, as explained in Fleiss et al. (2003) and applied by Kerr & Pender (2005) to farmers' perceptions about soil erosion in villages in India, to estimate the degree of agreement between farmers' perception about and measured soil fertility. The interrater agreement is measured using Cohen's kappa coefficient (Cohen, 1960). Kappa measures the degree of agreement beyond that which could be expected by chance (Fleiss et al. 2003). It measures the degree of agreement between raters in instances where the scales of ratings are categorical. To illustrate, consider dummy Table 1.1 below of data on agreement in ratings of soil fertility by farmers' perception and by measurement. Each of the cells, i.e. *a, b, c* and *d*, contains the proportion of all maize plots categorized as either fertile or infertile by farmers and fertile or infertile by measurement. For example, cell *a* contains the proportion of plots that are considered fertile according to both farmers' perception and measurement, while cell *b* contains the proportion of plots which the farmers assess as fertile but are measured to be infertile. The kappa statistic is expressed as in (2.3) and ranges from -1 to +1:

$$\hat{\kappa} = \frac{2(ad-b\ )}{p_1q_2+p_2q_1} \tag{2.3}$$

Table 2.1: Dummy table of data measuring agreement between farmers' perception and measured soil fertility

Rater A (Farmers' perception)	Rater B (Measured fertility)		Total
	Category 1 (Fertile)	Category 2 (Infertile)	
Category 1 (Fertile)	a	b	p_1
Category 2 (Infertile)	c	d	q_1
Total	p_2	q_2	1

Landis & Koch (1977) and Fleiss et al. (2003) have suggested the following interpretations of kappa ranges for degree of agreement:

Kappa ($\hat{\kappa}$) value range:	Degree of agreement beyond chance:
Landis & Koch (1977):	
$\hat{\kappa} < 0.00$	poor
$0.00 < \hat{\kappa} \leq 0.20$	slight
$0.20 < \hat{\kappa} \leq 0.40$	fair
$0.40 < \hat{\kappa} \leq 0.60$	moderate
$0.60 < \hat{\kappa} \leq 0.80$	substantial
$\hat{\kappa} > 0.80$	almost perfect
Fleiss et al. (2003):	
$\hat{\kappa} < 0.40$	poor
$0.40 < \hat{\kappa} \leq 0.75$	fair to good
$\hat{\kappa} > 0.75$	excellent

Although useful in gauging the extent of agreement between two raters, the suggested ranges of kappa and the agreement levels they represent are certainly arbitrary. We use both scales of agreement in interpreting our results.

Lastly, we apply a probit model to estimate the association between farmers' perception about and measured soil fertility as well as measured individual physical and chemical properties of soil. The aim is to explore whether there is a significant relationship between farmers' perceived and measured soil fertility and whether such a relationship holds when we account for the effects of other relevant factors. The model takes the following form:

$$Y = f(M, F), \qquad (2.4)$$

where M represents measured soil fertility and F represents other factors deemed to influence farmers' perceptions. Y is farmer's perception about soil fertility and is a binary variable for the probability of a farmer's perception of the soil on a plot as fertile. We define for farmer i a latent measure of perceived soil fertility on a plot as Y^*, such that

$$Y^*{}_i = \boldsymbol{Z_i\beta} + \varepsilon_i, \qquad (2.5)$$

where $i = 1, \dots., N$, $\boldsymbol{Z}$ is a row vector of explanatory variables (M, F) hypothesized to influence farmer's perception, $\boldsymbol{\beta}$ is a column vector of parameters to be estimated, and ε is a random error term, here assumed to be distributed normal with a constant variance (i.e. $\varepsilon \sim N(0, \sigma^2)$).

In this latent variable setting, we observe only Y_i (i.e. whether a farmer indicates that a plot is fertile or infertile):

$$Y_i = \begin{cases} 1\ if\ Y^*{}_i > 0\ (fertile) \\ 0\ otherwise\ (infertile) \end{cases}$$

$$Y^*_i > 0 \Rightarrow \boldsymbol{Z_i\beta} + \varepsilon_i > 0 \Rightarrow \varepsilon_i > -\boldsymbol{Z_i\beta} \qquad (2.6)$$

So, in the probit model (Wooldridge, 2010) $P(Y_i = 1|\boldsymbol{Z_i}) = P(Y^*_i > 0|\boldsymbol{Z_i}) = P(\varepsilon_i > -\boldsymbol{Z_i\beta}) = P\left(\frac{\varepsilon_i}{\sigma} > \frac{-\boldsymbol{Z_i\beta}}{\sigma}\right) = \Phi\left(\frac{-\boldsymbol{Z_i\beta}}{\sigma}\right) = G(\boldsymbol{Z_i\beta})$, where we set σ=1 and effectively make the distribution of the error term, ε, standard normal density. $G(.)$ is a cumulative normal distribution function for some continuous variable with a density $g(.)$. $P(Y_i = 0|\boldsymbol{Z_i}) = 1 - \Phi\left(\frac{-\boldsymbol{Z_i\beta}}{\sigma}\right) = 1 - G(\boldsymbol{Z_i\beta})$ is similarly derived. We obtain the estimates of $\boldsymbol{\beta}$ using maximum likelihood estimation procedure.

The average partial effect of a continuous explanatory variable *j*, on the response probability (in this case the probability of a farmer indicating that the soil on a plot is fertile) is derived as:

$$\widehat{APE_j} = \widehat{\beta_j}\left[N^{-1}\textstyle\sum_i^N g(\boldsymbol{Z_i}\widehat{\boldsymbol{\beta}})\right] \qquad (2.7)$$

For a binary explanatory variable *k*, the average partial effect on the response probability is derived as follows:

$$\widehat{APE_k} = N^{-1}\textstyle\sum_i^N\left[G\left(\boldsymbol{Z_{i(k)}}\widehat{\boldsymbol{\beta}_{(k)}} + \widehat{\beta_k}\right) - G\left(\boldsymbol{Z_{i(k)}}\widehat{\boldsymbol{\beta}_{(k)}}\right)\right], \qquad (2.8)$$

where $\boldsymbol{Z_{i(k)}}$is $\boldsymbol{Z_i}$ excluding the explanatory variable *k*.

We augment these analyses with mean comparison test of a range of soil physical and chemical properties, plot attributes, input use, agronomic management practices, plot manager attributes, and maize yield between farmers grouped according to their perceptions about soil fertility.

2.2.2.2 Adoption of soil fertility management practices

A number of studies on adoption of soil fertility management practices have found simultaneity of adoption decisions of different practices by farmers (e.g. Marenya & Barrett, 2007; Kamau et al., 2014; Kassie et al., 2015; Wainaina et al., 2016; Koppmair et al., 2017). Failure to account for such simultaneity in econometric analysis of adoption decisions may compromise efficiency of the estimates and make inferences inaccurate. Therefore, this analysis applies a multivariate probit model because there are multiple soil fertility management practices for which adoption variable is binary and farmers' adoption decisions may be jointly made.

Multivariate probit model is like a seemingly unrelated regression for binary response variables. We express the system of equations to be estimated as a series of equation (2.2) that match the number of soil fertility management practices:

$$T_{i1}^{*} = \boldsymbol{V}_{i1}\boldsymbol{\beta}_{\mathbf{1}} + \varepsilon_{i1}$$

$$T_{i2}^{*} = \boldsymbol{V}_{i2}\boldsymbol{\beta}_{\mathbf{2}} + \varepsilon_{i2}$$

$$\vdots$$

$$T_{iG}^{*} = \boldsymbol{V}_{iG}\boldsymbol{\beta}_{\boldsymbol{G}} + \varepsilon_{iG}\ , \qquad (2.9)$$

with $T_{ig} = 1[T_{ig}^{*} > 0]$ and $\boldsymbol{\varepsilon}_i|\boldsymbol{V}_i \sim Normal\ (\mathbf{0}, \boldsymbol{\Omega})$ with unit variances. This means that the marginal distributions conditional on $\boldsymbol{V}$ are assumed to follow probit:

$$P(T_g = 1|\boldsymbol{V}) = \boldsymbol{\Phi}\left(\boldsymbol{X}_g\boldsymbol{\beta}_g\right), \qquad g = 1, \dots, G$$

2.2.3 Data sources and variables

2.2.3.1 Data sources

The study uses household- and plot-level survey data on maize production in Kenya. The data were collected by the Department of Agricultural, Food and Resource Economics at Michigan State University under the Guiding Investments in Sustainable Agricultural Intensification in

Africa (GISAIA) research project implemented in seven countries in Africa. The project aimed to influence policy environment to improve sustainable intensification of cereal crop production and contribute to sustainable growth of agricultural productivity in Africa. In Kenya, the project activities focused on efficiency and profitability of use of inorganic fertilizer and hybrid seed, and agricultural land constraints and changing farm structure.

The first survey was conducted in 2014 from a sample of 650 farm households spread in five counties. Sample selection followed a multistage procedure. Five counties were purposively selected in the first stage based on two criteria: counties that are most important in maize production to study maize yield response to inorganic fertilizer use; and highly densely populated counties to study population density-intensification relationship. Uasin Gishu, Kakamega and Trans Nzoia counties were selected for maize yield response study while Machakos and Kisii counties were selected for population density-intensification relationship study. In the second stage, two sub-counties were selected in Uasin Gishu, Kakamega and Trans Nzoia counties, with one in each representing a locality perceived to be of low and the other of high maize yield response to fertilizer. In Machakos and Kisii, two and three sub-counties, respectively, with highest population densities within the county were selected. In the third and subsequent stages, Locations, Sub-locations, villages and farm households in that order were randomly selected. The sample distribution across the administrative units is shown in Table 2.2.

The second survey was conducted in 2016 on the same households as in the first survey, but the number reduced to 623 because some of the households could not be reached for re-interview for various reasons, including migration. This survey, however, did not necessarily target the same maize plots that were cultivated in the 2014 survey. This means that the household-level data is panel while plot-level data is not necessarily panel. While there was a question in the 2016 survey questionnaire asking whether the largest maize plot (which is used in this study) was the same as in the 2014 survey, the responses in the affirmative were not consistent with plot sizes compared between 2016 and 2014. For that reason and because plot-level data is central in this study, the data from the two surveys are treated as pooled cross sections rather than panel.

Plot-level data includes crop output, types and amounts of inputs used, soil management practices, farmer's perception about soil fertility, and soil physical and chemical properties and were obtained for the largest maize plot cultivated by each household in the main season of 2013/2014 and 2015/2016 cropping years[2]. Willy, Muyanga, & Jayne (2016) document details about soil sampling and testing procedures. On average, the share of the largest maize plot in a household's total area of plots that had maize in them in the main season was 83%, indicating that this analysis captures the vast majority of maize produced by the farm households in the sample.

Table 2.2: Distribution of sample households across administrative units, 2014

County	Sub-county	Number of Sub-locations	Number of Villages	Number of households
Uasin Gishu	Wareng	2	8	60
	Eldoret West	2	5	60
Trans Nzoia	Kwanza	1	4	60
	Saboti	2	4	60
Kakamega	Kakamega North	2	4	60
	Lugari	1	4	60
Kisii	Bobasi	2	5	60
	Marani	2	5	60
Machakos	Kangundo	1	5	85
	Kathiani	1	3	51
	Machakos	1	2	34
		17	**49**	**650**

[2] Agricultural production data collected in 2014 and 2016 pertain to 2013/2014 and 2015/2016 cropping year, respectively.

2.2.3.2 Variables

Farmers' perceived and measured soil fertility

Farmers' beliefs and/or knowledge about the fertility condition of soil on their farms can partly influence how they manage their agricultural production on the farm. However, such beliefs may not necessarily reflect the actual soil fertility condition and thus may result in management practices that mismatch the fertility needs of the soil. Farmers were asked to rate the fertility of soil on their maize plots on a Likert scale of 1 (very infertile) to 4 (very fertile) based on their own perceptions. Because of very low responses on the two extreme categories (1 and 4) we reduce the scale to two categories - infertile (combining 1 and 2) and fertile (combining 3 and 4) to generate a binary variable of farmer perceived soil fertility.

From the soil test data, we measure the fertility of soil on a plot using three chemical properties of soil – total carbon (C), total nitrogen (N) and pH. IIASA/FAO (2012) suggest soil nutrient availability and retention capacity as among seven soil qualities that affect crop performance. Others are oxygen availability to roots, rooting conditions, toxicities, salinity and sodicity and workability. The authors suggest that natural nutrient availability in the soil is essential for low to medium input cropping, while nutrient retention capacity is especially important for effectiveness of inorganic fertilizer use. Therefore, these two soil qualities are most relevant in the context of maize production in Kenya where external input use is not that high and inorganic fertilizer use appears not to generate significant yield gains. As explained earlier, soil organic carbon influences a range of soil properties, including soil moisture and nutrient holding capacities, microbiological activity, and soil structure and other physical properties and is often used as an indicator for soil health (Okalebo, Gathua, & Woomer, 2002; Lal, 2006; Horneck et al., 2011). Soil pH influences nutrient availability and serves as an indicator for micro-nutrient deficiencies (IIASA/FAO, 2012). Together, IIASA/FAO (2012) suggest, soil organic carbon and pH are the best simple gauge for soil health. Nitrogen is the nutrient required by plants in largest amounts and is often the limiting nutrient in the soil.

We use threshold values for the properties as obtained from recommendations by the Kenya government on critical levels of various soil nutrients and pH for maize growth (Government of Kenya, 2014). A plot is measured as fertile if the soil has C$\geq 2.7\%$, N$\geq 0.2\%$ and $5.5 \leq \text{pH} \leq 7.0$ and infertile if it does not meet at least one of these three thresholds. We compare the farmer's subjective rating (fertile or infertile) against the measured fertility rating (fertile or infertile) as per these threshold values.

Correlates of farmer perception about soil fertility

For the probit model of the relationship between farmers' perception about soil fertility and individual soil physical and chemical properties, we use as explanatory variables a range of measured soil properties. In addition, we control for other relevant plot and plot manager characteristics. Texture is the only information about soil physical properties available in the data. We classify soil texture based on relative percentages of sand, silt and clay and according to the classification in (IIASA/FAO, 2012). Thus, we have four texture classes; sandy, loamy coarse, loamy moderate, loamy fine and clayey soils. Among the soil chemical properties proposed by Doran & Parkin (1994) as indicators that can be used to assess soil fertility include total organic carbon, total nitrogen, soil pH, electrical conductivity and extractable (or plant-available) macronutrients (P, K). Because total carbon and total nitrogen are highly correlated, we excluded total nitrogen in the estimation. Soil pH measures acidity or alkalinity and its level affects biological and chemical activity in the soil and affects availability of nutrients to plants. Lower pH levels indicate greater acidity while higher values indicate alkalinity. Most crops require near neutral levels of pH to grow best. Soil electrical conductivity is a measure of the amount of soluble (salt) ions in the soil and is strongly correlated with many soil properties, particularly texture (Grisso et al, 2009). We thus do not include it in the model. We include plant-available phosphorus among the macronutrients because it is particularly important for crop growth and is sensitive to soil acidity level.

The plot characteristics we include are farmer-reported slope, measured in categories – flat, moderate and steep, and the number of years during the last decade the plot was in cultivation. We include distance (in walking minutes) of plot from the homestead and plot manager's education, farming experience and gender to capture the effects of plot manager's characteristics. To test whether perception is correlated with information access, we include variables that measure area physical infrastructure – distance to extension service provider, distance to town and distance to paved road.

Determinants of adoption of soil fertility management practices

We consider five soil fertility management practices in the adoption analysis: inorganic fertilizer, manure/compost, crops residue, legume intercrop and soil erosion control. Adoption of each of the practices is measured as a binary variable, taking the value of 1 if a practice was implemented on a plot and 0 if not. It is important to recognize that the variable for soil erosion control is not really a measure of whether the farmer invested in soil erosion control during the cropping season but whether there were soil erosion structures on the plot irrespective of when they were established. This is because measures for soil erosion control such as terraces, grass strips and cut-off drains tend to be relatively permanent, with regular or occasional maintenance. However, the data does not have information about whether the farmer carried out any maintenance on the soil erosion structures on the plot.

Adoption rates of the practices are presented in Table 2.3. Inorganic fertilizer is the most important external inputs into maize production and which is accorded much attention by both the government and farmers in efforts to raise maize productivity. It is important to note that adoption rate of inorganic fertilizer is that high because a large part of the sample is concentrated in the major maize growing regions – Uasin Gishu, Trans Nzoia and Kakamega. Because of their role as the country's grain basket areas, these regions have had a history of concerted promotion by both the government and development agencies of adoption of fertilizers and improved maize varieties. Manure and compost, in addition to supplying nutrients into the soil, increase soil organic matter, and therefore organic carbon, which is beneficial for soil fertility as explained in the introduction. Crops residue also add organic matter in the soil, and if used as mulch conserve moisture in the soil and protect the topsoil from erosion that may result from direct impact of rain. Intercropping maize with legumes, such as common beans, groundnuts and pigeon peas, enrich the soil with nitrogen because they can fix atmospheric nitrogen into the soil. Soil erosion control, through structures such as terraces, grass strips and cut-off drains, is beneficial in preventing soil loss and loss of soil nutrients with it. It is our hypothesis that farmers may make joint decision regarding adoption of some of these practices hence the need for simultaneous estimation of adoption equations to ensure efficiency of estimated parameters.

Table 2.3: Adoption rates of various soil fertility management practices (N=1268)

Management practice	Mean	SD
Inorganic fertilizer (1=yes, 0 otherwise)	0.93	0.26
Manure/compost (1=yes, 0 otherwise)	0.46	0.50
Crops residue (1=yes, 0 otherwise)	0.41	0.49
Legume intercrop (1=yes, 0 otherwise)	0.71	0.45
Soil erosion control (1=yes, 0 otherwise)	0.51	0.50

Literature on adoption of soil fertility management practices have found a range of factors that influences adoption (e.g. Marenya & Barrett, 2007; Odendo et al., 2009; Kamau et al., 2014; Kassie et al., 2015; Wainaina et al., 2016; Arslan et al., 2017; Koppmair et al., 2017). We present the descriptive statistics of the factors included in this analysis in Table 2.4. The factor of interest in this study is farmer's soil fertility perception, which most adoption studies do not include among explanatory variables that explain technology adoption decisions. We include plot characteristics among the factors since the condition of a plot, such as slope, may make particular practices better suited to and necessary on it. The characteristics we include are slope and number of years within the last decade the plot has been on cultivation. Plot manager characteristics include walking time to the plot, whether the plot is owned, which is a measure of tenure security, and plot manager's education, gender and farming experience.

Three measures of household capital are included. Number of adults in a household represents human capital in terms of labor, while group membership represents social capital. Group membership may also act as a mechanism for access to information and credit, both of which have been found to be important to technology adoption. Whether a household has non-farm income may be an indicator of working capital, which is important for acquisition of external inputs such as inorganic fertilizers and hired labor. The ability to hire labor may be important to adoption of labor-intensive soil management practices such as manure/compost and crops residue use.

Table 2.4: Summary statistics of explanatory variables in the adoption estimation (N=1268)

Explanatory variables	Mean	SD
Soil fertility perception (1=fertile, 0 otherwise)	0.59	0.49
Years of plot cultivation	9.56	1.55
Plot slope (1=moderate, 0 otherwise)[a]	0.54	0.50
Plot slope (1=steep, 0 otherwise)[a]	0.10	0.30
Plot size (acres)	1.11	2.19
Walking time to plot (mins)	4.73	10.85
Plot owned (1=yes, 0 otherwise)	0.90	0.30
Plot manager education (yrs)	8.02	3.86
Plot manager is female (1=yes, 0 otherwise)	0.35	0.48
Farming experience (yrs)	27.52	16.22
Active adults (#)	2.99	1.75
Group membership (1=yes, 0 otherwise)	0.72	0.45
HH has non-farm income (1=yes, 0 otherwise)	0.72	0.45
Livestock value (KES) (log)	9.72	3.13
HH landholding (acres)	2.47	25.81
Asset index	0.02	0.07
Distance to extension (km)	5.69	5.07
Distance to paved road (km)	6.92	6.87
Fertilizer price (KES/kg)	64.31	18.09

[a] Base category: flat

We measure household wealth using value of livestock, landholding and asset index. On one hand, livestock are important because they produce manure and, therefore, may motivate manure use. On the other hand, ruminants are often fed crops residue and thus presence of livestock may have a negative effect on the adoption of the practice. Size of landholding may determine a household's motivation for agricultural intensification, with those having smaller landholdings having higher propensity to apply intensification practices in a bid to maximize yield. Asset index, and indicator of the capacity for access to resources, such as credit, is computed from a list of household assets (agricultural and household) using principal component analysis. The index is based on the score for the first principal component, which accounts for the largest variation in the data.

Information and market access conditions are proxied by distance to extension information and distance to paved road. These variables are generally expected to have negative effects on adoption of soil fertility management practices. Fertilizer price is included and is expected to be negatively correlated with adoption of fertilizer and practices for which adoption is complementary to that of fertilizer.

2.3 Results

2.3.1 Farmers' perceived versus measured soil fertility

As explained in the previous section, farmers were asked to rate the fertility of soil on their largest maize plots on a Likert scale of 1 (very infertile) to 4 (very fertile) based on their own perceptions. Because of very low responses on the two extreme categories (6% for very infertile and 5% for very fertile), we reduced the scale to two categories - infertile (combining 1 and 2) and fertile (combining 3 and 4) thus resulting into a binary variable indicating farmers' perception about the fertility status of their plots as either fertile or infertile. Results show that overall, 59% of the plots were perceived to be fertile while 41% were perceived to be infertile. Across regions (counties), higher percentages of plots in Kakamega and Kisii were perceived to be infertile compared to those in the other counties (Figure 2.2).

Using data from the laboratory test of physical and chemical properties of soil from the same plots of which farmers subjectively rated fertility, we constructed a measure of fertility for each plot. As explained in the preceding section, a plot was measured to be fertile if the soil had total C$\geq 2.7\%$, N$\geq 0.2\%$ and $5.5 \leq \text{pH} \leq 7.0$ and infertile if at least one of these conditions was not met. Based on this index, results show that 14% of the plots were measured to be fertile while 86% were infertile. Figure 2.3 shows that incidence of soil infertility according to this measure was highest in Kakamega, Machakos and Kisii in that order and lowest in Trans Nzoia and Uasin Gishu. It is worth noting that population densities are higher in Machakos, Kisii and Kakamega than in Uasin Gishu and Trans Nzoia, indicating that soil infertility problem could be worse in areas where land scarcity is more acute.

Overall, farmer-perceived and measured fertility status of 46% of plots were consistent, about half of the plots were rated as fertile by farmers but measured to be infertile, and less than 5% of the plots were perceived as infertile but measured to be fertile.

Figure 2.2: Percentages of farmer-perceived fertile and infertile plots across regions

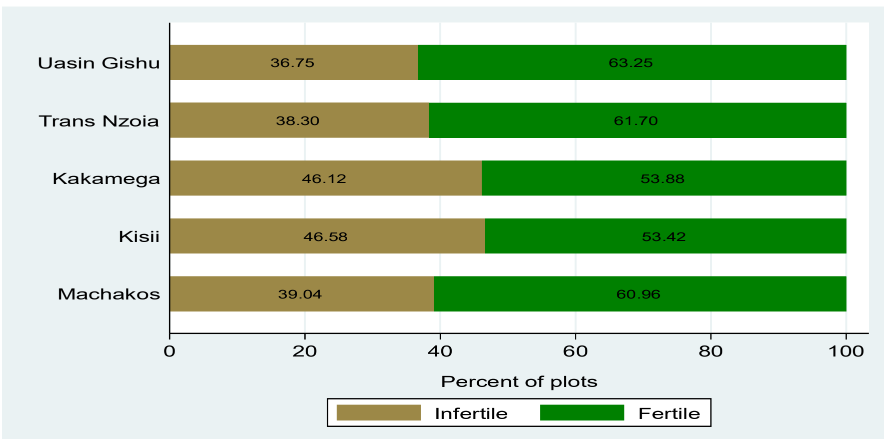
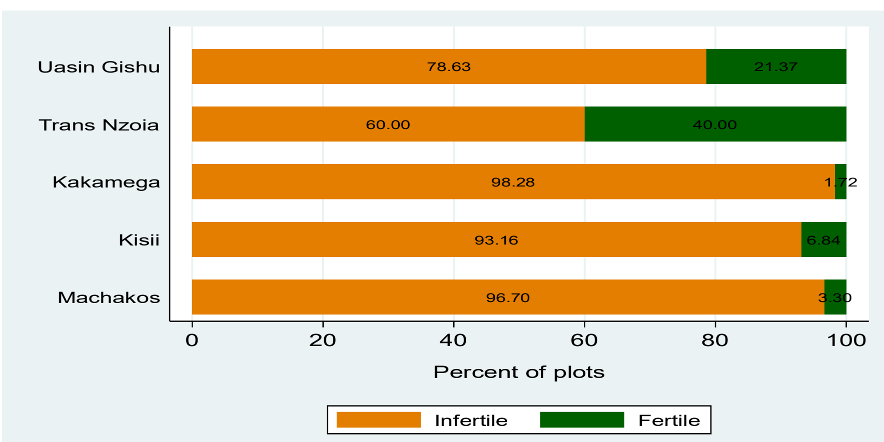


Figure 2.3: Percentages of measured fertile and infertile plots across regions



Using the chi-square test of independence between farmers' perceived and measured soil fertility, results show that overall, the two soil fertility ratings are not statistically independent at the 95% confidence level (Table 2.5). However, the analysis separated by region rejects (at the 95% confidence level) the null hypothesis of independence only in Machakos county. This suggests that perceived and measured soil fertility are largely misaligned for majority of the plots, reflecting the earlier description that 54% of the plots had farmers' perceived inconsistent with measured soil fertility.

The test results for the extent of agreement between farmers' perceived and measured soil fertility using the interrater agreement technique is presented in Table 2.6. Using the Landis & Koch (1977) and Fleiss et al. (2003) scales of degree of agreement, the overall kappa value of 0.037, although statistically significant, indicates only slight (Landis & Koch, 1977) or poor (Fleiss et al., 2003) agreement beyond that which could be expected by chance between farmer perceived and measured soil fertility. This conclusion also generally obtains in each of the five regions.

Table 2.5: Chi-square test of independence between farmer perceived and measured soil fertility

Overall sample			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	463	59	522
Fertile	630	116	746
Total	1093	175	1268
Pearson chi2(1) = 4.656 Pr = 0.031 ; Fisher's exact = 0.032			
Uasin Gishu			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	66	20	86
Fertile	118	30	148
Total	184	50	234
Pearson chi2(1) = 0.289 Pr = 0.591; Fisher's exact = 0.622			
Trans Nzoia			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	60	30	90
Fertile	81	64	145
Total	141	94	235
Pearson chi2(1) = 2.701 Pr = 0.100; Fisher's exact = 0.132			
Kakamega			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	105	2	107
Fertile	123	2	125
Total	228	4	232
Pearson chi2(1) = 0.0246 Pr = 0.875; Fisher's exact = 1.000			
Kisii			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	103	6	109
Fertile	115	10	125
Total	218	16	234
Pearson chi2(1) = 0.569 Pr = 0.451; Fisher's exact = 0.605			
Machakos			
Farmer perceived fertility	Measured fertility		
	Infertile	*Fertile*	*Total*
Infertile	129	1	130
Fertile	193	10	203
Total	322	11	333
Pearson chi2(1) = 4.287 Pr = 0.038; Fisher's exact = 0.056			

Table 2.6: Kappa statistic measuring degree of agreement between perceived and measured soil fertility

County	% agreement	% expected agreement	Kappa	Std. Error	Z	Prob>Z
Uasin Gishu	41.03	42.41	-0.024	0.045	-0.54	0.7044
Trans Nzoia	52.77	47.66	0.098	0.059	1.64	0.0501
Kakamega	46.12	46.25	-0.003	0.016	-0.16	0.5624
Kisii	48.29	47.05	0.024	0.031	0.75	0.2253
Machakos	41.74	39.76	0.033	0.016	2.07	0.0192
Overall sample	45.66	43.61	0.037	0.017	2.16	0.0155

It is useful to understand how farmers' perceived soil fertility correlates with measured fertility, measured individual physical and chemical soil properties and other plot and farm characteristics. Such information is helpful in identifying factors that contribute to heterogeneity in individual farmers' perceptions about soil fertility. As explained earlier, we explore this by estimating a probit model where the dependent variable is the binary indicator of farmer's perception about soil fertility, and explanatory variables are measured soil fertility, selected soil physical and chemical properties, other plot, plot-manager and area characteristics and maize yield. Before discussing the probit model estimation results, we explore difference in means of a range of variables between plots farmers perceived to be fertile and those perceived as infertile (Table 2.7).

On average, plots farmers perceived to be fertile were significantly richer in organic carbon and nitrogen, had higher concentration of potassium and were less acidic (had higher soil pH) (Table 2.7). Perception about soil fertility, however, appears not to be associated with soil texture in a significant way, as can be seen in the difference in the means of sand, silt and clay content between perceived fertile and infertile plots. These statistics suggest that on average, farmers' soil fertility perception have some alignment with levels of individual soil chemical properties that influence soil nutrient availability and retention capacity, and hence fertility. We revisit this conjecture when discussing the probit model estimation results.

Table 2.7: Mean comparison test of selected variables according to farmers' perception about soil fertility status

Attributes	Farmer perception of soil fertility status		Difference in mean (F-I)
	Fertile (F)	Infertile (I)	
Plot characteristics			
Total organic carbon (%)	2.180	2.050	0.130**
Total Nitrogen %	0.188	0.178	0.00935**
Phosphorus (ppm)	19.776	17.870	1.906
Potassium (meq %)	0.820	0.748	0.0718*
Soil pH	5.727	5.665	0.0626**
Sand (%)	48.879	49.171	-0.293
Silt (%)	12.244	12.482	-0.238
Clay (%)	38.745	38.137	0.609
No. of years of plot cultivation in the past 10 years	9.507	9.638	-0.131
Plot slope (1=flat, 2=moderate, 3=steep)	1.728	1.770	-0.0422
Input use on plot			
Inorganic fertilizer use (1=yes, 0 otherwise)	0.913	0.948	-0.0354**
Fertilizer rate (kg/acre) [users only]	81.723	80.741	0.982
Improved seed variety (1=yes, 0 otherwise)	0.928	0.927	0.000411
Mechanized land preparation (1=yes, 0 otherwise)	0.532	0.450	0.0820***
Agronomic practices on plot			
Manure/compost use (1=yes, 0 otherwise)	0.472	0.454	0.0178
Crops residue use (1=yes, 0 otherwise)	0.401	0.427	-0.0264
Maize-legume intercrop (1=yes, 0 otherwise)	0.685	0.757	-0.0717***
Soil erosion prevention (1=yes, 0 otherwise)	0.504	0.519	-0.0151
Maize production			
Maize yield (kg/acre)	1398.683	1183.586	215.1***
Local infrastructure			
Distance to extension (km)	5.434	6.059	-0.625**
Distance to town (km)	13.460	13.868	-0.408
Distance to paved road (km)	7.009	6.783	0.226
Plot manager characteristics			
Plot manager education (yrs)	8.059	7.962	0.0973
Plot manager is female (1=yes, 0 otherwise)	0.336	0.358	-0.0218
Farming experience (yrs)	27.657	27.316	0.341
Walking time to plot (mins)	5.288	3.925	1.362**
Plot owned (1=yes)	0.883	0.916	-0.0323*
Observations	746	522	1268

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Plots perceived to be fertile were on average located further from the homestead than those perceived to be infertile. In addition, incidence of ownership (rather than renting) was lower for plots perceived to be fertile than for those perceived to be infertile. This indicates that rented plots, which the data show are on average located farther from the homestead than are owned plots, are on average more likely to be rated fertile. It implies that, as would be expected, farmers who rent plots for maize production have a higher likelihood to go for those they perceive to be fertile.

While inorganic fertilizer use was generally high (about 93% of the plots in total), results show some association between the binary decision to use inorganic fertilizer and perception about soil fertility: inorganic fertilizer use was more likely on plots perceived to be infertile than on those perceived to be fertile. Whether this reflects conscious efforts by farmers to address perceived soil infertility through use of inorganic fertilizer is a question deserving investigation. We explore the possibility of this in the discussion of results on adoption of soil fertility management practices. Conditional on use, however, mean application rates of inorganic fertilizer were statistically the same between the two groups of plots. Incidence of mechanized land preparation was higher on plots that were perceived to be fertile. Among agronomic practices, we only observe significantly higher incidence of maize-legume intercrop on plots perceived to be infertile than on those perceived to be fertile, while with regards to area infrastructure mean distance from extension advice was shorter for plots that were perceived to be fertile than those that were perceived to be infertile. As would be expected, the mean yield of maize was significantly higher on plots perceived to be fertile than on those perceived to be infertile.

In summary, plots farmers perceived as fertile relative to those perceived to be infertile had, on average, higher levels of individual soil chemical attributes that influence nutrient availability and were less likely to have inorganic fertilizer applied. In addition, they were more likely to be intercropped with legumes and had higher maize yield. Further, such plots were located farther away from the homestead, suggesting that farmers viewed rented plots as more fertile.

Estimated marginal effects from a sequence of probit models of farmers' perception about soil fertility are presented in Tables 2.8 and 2.9. In Table 2.8 are the marginal effects of measured soil fertility, along with other covariates, while in Table 2.9 are marginal effects of individual soil properties, along with other covariates. The tables each has a sequence of five models to show the behavior of the effects of measured soil fertility and individual soil properties when other covariates are included. The other covariates are plot characteristics, maize yield, plot manager characteristics, and area infrastructure. In all the specifications, regional effects are controlled for using regional (county) dummy variables. A squared term for the soil pH variable was included in the estimation of models in Table 2.9 to account for nonlinearity of soil pH, since in actual sense extreme values of soil pH are associated with suboptimal conditions for plant growth.

In Table 2.8, model 1 shows that there is a positive but statistically weak correlation between farmers' perceived and measured soil fertility status, holding other things constant. Compared to a plot measured to be infertile, a plot that is measured to be fertile has, on average, 7% higher chance of being rated as fertile by a farmer. The small magnitude of the marginal effect of measured soil fertility combined with its weak statistical significance strengthens the findings from the interrater agreement through the Cohen's kappa statistic and the chi-square test of independence between farmers' perceived and measured soil fertility. When we control for maize yield, however, the weak significance of the association between farmers' perceived and measured soil fertility vanishes and the marginal effect of maize yield is positive and strongly significant (model 2), a result that also obtains in models 3, 4 and 5, which have plot characteristics, plot-manager characteristics and area infrastructure, respectively. The negative and significant marginal effects of steep slope dummy variable in models 3-5 is evidence that plot slope is significantly associated with farmers' perception about soil fertility. In addition, the negative and significant marginal effect of distance to extension service provider indicates that access to extension information may have influence on farmers' perception about the fertility of their soils.

Table 2.8: Probit marginal effect estimates of the relationship between farmers' perceived and measured soil fertility

Dep var: Soil fertility perception (1=fertile)	(1) *Measured fertility*	(2) *Maize yield*	(3) *Plot characteristics*	(4) *Plot manager characteristics*	(5) *Area infrastructure*
Measured fertility (1=fertile)	0.0732*	0.0707	0.0711	0.0698	0.0719
	(0.0443)	(0.0441)	(0.0444)	(0.0443)	(0.0456)
Maize yield (kg/acre)		0.0000513***	0.0000512***	0.0000510***	0.0000497***
		(0.0000146)	(0.0000146)	(0.0000147)	(0.0000147)
Moderate[b]			-0.0252	-0.0230	-0.0243
			(0.0309)	(0.0310)	(0.0310)
Steep[b]			-0.104*	-0.111**	-0.0991*
			(0.0556)	(0.0557)	(0.0562)
No. of years of plot cultivation			-0.0107	-0.00630	-0.00673
			(0.00917)	(0.00955)	(0.00953)
Plot manager education (yrs)				0.000548	0.000563
				(0.00399)	(0.00398)
Plot manager is female (1=yes)				-0.0168	-0.0168
				(0.0301)	(0.0301)
Farming experience (yrs)				0.000749	0.000667
				(0.000948)	(0.000950)
Plot owned (1=yes)				-0.0558	-0.0619
				(0.0487)	(0.0487)
Walking time to plot (mins)				0.00256	0.00241
				(0.00158)	(0.00158)

Table 2.8 (cont'd)

Dep var: Soil fertility perception (1=fertile)	(1) *Measured fertility*	(2) *Maize yield*	(3) *Plot characteristics*	(4) *Plot manager characteristics*	(5) *Area infrastructure*
Distance to extension (km)					-0.00473*
					(0.00276)
Distance to paved road (km)					-0.000476
					(0.00298)
Regional dummies	Yes	Yes	Yes	Yes	Yes
Observations	1268	1268	1268	1268	1268
LR χ^2	11.19	23.35	28.22	34.19	38.72
Prob > χ^2	0.0477	0.0007	0.0009	0.0019	0.0020
Pseudo R^2	0.0065	0.0136	0.0164	0.0199	0.0225

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results in Table 2.9 are largely similar to those in Table 2.8. In model 1 of Table 2.9 where the only covariates – apart from regional dummies – are individual soil properties, the marginal effects of soil organic carbon and pH are positive and significant (albeit weakly). A unit increase in soil pH is associated with 5% higher chance of the plot being rated as fertile by a farmer, while a one-percentage point increase in soil organic carbon content is associated with 3% higher chance of a plot being rated as fertile by a farmer. These results strengthen the findings in Table 4 that farmers' soil fertility perceptions have some reflection of levels of individual soil chemical properties that influence soil nutrient availability. However, the statistical significance of the marginal effects of soil organic carbon vanishes when we introduce other covariates in models 2-5, while that of soil pH remains significant only in model 4. Like in Table 8, the marginal effect of maize yield remains positive and significant in models 2-5 and that of the steep slope dummy is negative and significant in models 3-5.

The evidence that the marginal effects of measured soil fertility (Table 2.8) and those of soil organic carbon and pH (Table 2.9) become insignificant when other covariates are introduced is evidence that farmers' perceptions about soil fertility are based on factors other than measurable soil attributes. The important factors that appear to influence farmers' perception about soil fertility are maize yield and plot slope, with maize yield showing a strong statistical correlation in all the models in which it is included. This finding is consistent with that of Marenya et al (2008), who suggested that famers may be prone to error in their soil fertility assessment given the many factors, in addition to soil conditions, that influence yield. When observed yield levels is the primary information on which farmers base their assessment of soil fertility status, they are likely to apply soil fertility management practices that may not match the soil fertility needs and this can lead to further soil degradation. The data in this research show that 39% of the plots in the sample had soil pH values below the lower bound of the recommended optimal range of 5.5-7.0 while 52% had values below the sample average of 5.7, which is barely above the lower bound of the range. In the contrary, only 35% of the farmers acknowledged that their plots had soil acidity problems, 38% indicated there were no soil acidity problems on their plots, while the remaining 27% had no idea regarding whether their plots had soil acidity problems. Therefore, it is not surprising that majority of the farmers in the sample apply DAP, an acidifying fertilizer, in efforts to raise maize yield.

Figure 2.4 shows that among the plots that had soil pH<5.5 (39% of plots), 44% had DAP applied without manure or compost while 11% had neither DAP nor manure/compost application. About 35% had both DAP and manure/compost while 12% had manure/compost without DAP. A similar pattern is observed on plots that had low organic carbon (<2.7%) (75% of plots), where majority (44%) had DAP applied without manure/compost while 9% had neither DAP nor manure application (Figure 2.5). About one-third of the plots had both DAP and manure/compost applied while 17% had manure/compost only. The scenario in Figure 2.6 is most illuminating. Out of the plots that had both low soil pH (<5.5) and low organic carbon (<2.7%), 46% had DAP applied without manure/compost. These statistics show that even on plots that are low in soil pH and organic carbon, and on which less-acidifying fertilizers and organic soil amendments should be most appropriate, use of DAP without manure/compost is still the most common. But surprisingly, despite producing a report that shows that soil acidity and low soil organic carbon levels are widespread on farms in the major maize growing areas and providing recommendations for appropriate types and amounts of fertilizer and manure to apply (Government of Kenya, 2014), the government continues to subsidize DAP for maize production. These indicate deficiencies in extension information and policy and exacerbates poor management of soil fertility.

Figure 2.4: Frequency distribution of DAP and manure use on plots with low soil pH (<5.5) (39% of plots)

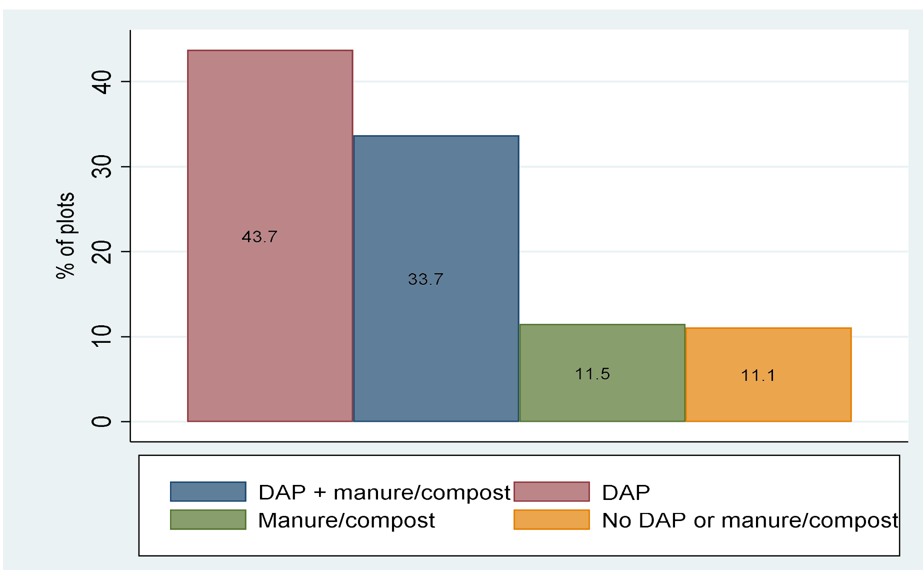


Figure 2.5: Frequency distribution of DAP and manure use on plots with low organic carbon (<2.7%) (75% of plots)

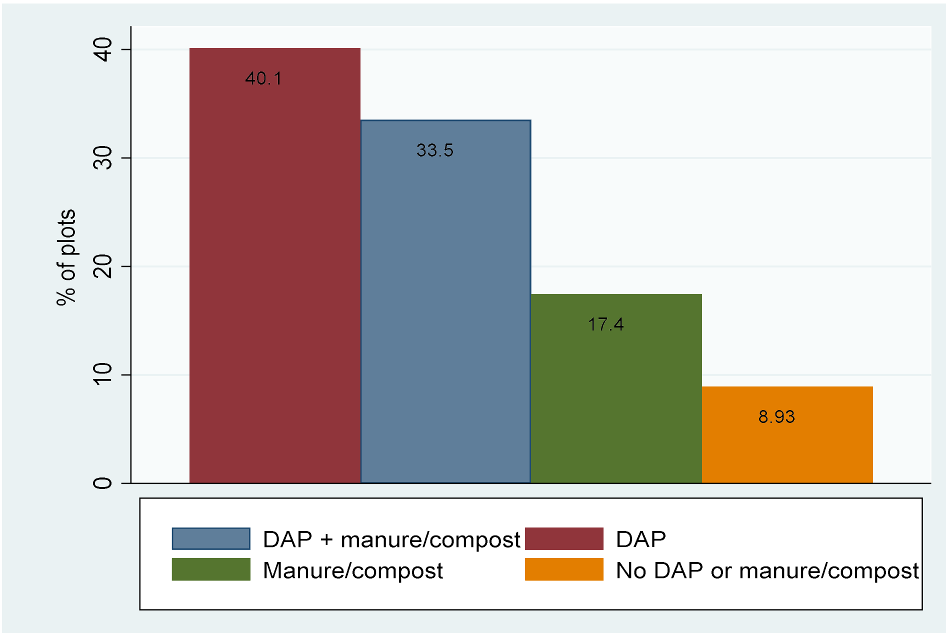


Figure 2.6: Frequency distribution of DAP and manure use on plots with both low organic carbon (<2.7%) and low soil pH (30% of plots)

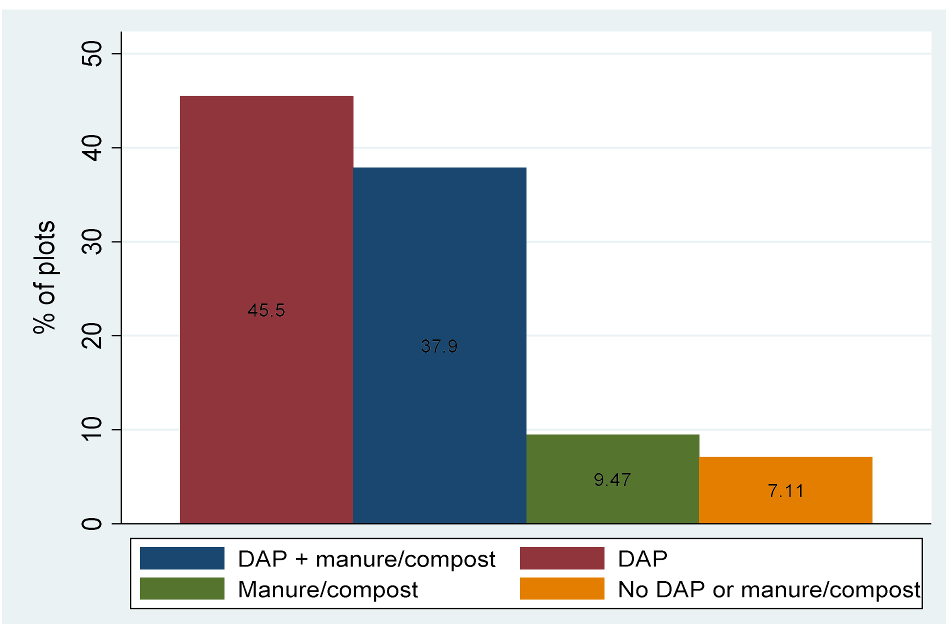


Table 2.9: Probit marginal effect estimates of the relationship between farmers' perception about soil fertility and measured soil properties

Dep var: Soil fertility perception (1=fertile)	(1) *Individual soil properties*	(2) *Maize yield*	(3) *Plot characteristics*	(4) *Plot manager characteristics*	(5) *Area infrastructure*
Total carbon (%)	0.0325*	0.0297	0.0305	0.0292	0.0305
	(0.0196)	(0.0196)	(0.0198)	(0.0198)	(0.0202)
pH (value)	0.0506*	0.0445	0.0472	0.0496*	0.0448
	(0.0298)	(0.0298)	(0.0298)	(0.0301)	(0.0303)
Phosphorus (ppm)	0.000783	0.000775	0.000589	0.000531	0.000554
	(0.000752)	(0.000746)	(0.000752)	(0.000752)	(0.000752)
Loamy coarse[a]	-0.0976	-0.0960	-0.0971	-0.0812	-0.0791
	(0.0971)	(0.0970)	(0.0964)	(0.0978)	(0.0974)
Loamy medium[a]	-0.172	-0.158	-0.176	-0.160	-0.159
	(0.190)	(0.190)	(0.188)	(0.188)	(0.187)
Loamy fine[a]	-0.0160	-0.0149	-0.0119	-0.000341	-0.00123
	(0.0910)	(0.0910)	(0.0903)	(0.0916)	(0.0912)
Clayey[a]	-0.0329	-0.0346	-0.0350	-0.0193	-0.0223
	(0.0926)	(0.0925)	(0.0919)	(0.0932)	(0.0928)
Maize yield (kg/acre)		0.0000484***	0.0000484***	0.0000477***	0.0000467***
		(0.0000146)	(0.0000146)	(0.0000147)	(0.0000147)
Moderate[b]			-0.0355	-0.0331	-0.0338
			(0.0313)	(0.0313)	(0.0313)
Steep[b]			-0.109*	-0.115**	-0.104*
			(0.0564)	(0.0565)	(0.0570)
No. of years of plot cultivation			-0.0102	-0.00570	-0.00609
			(0.00917)	(0.00954)	(0.00953)
Plot manager education (yrs)				0.000819	0.000818
				(0.00399)	(0.00398)
Plot manager is female (1=yes)				-0.0161	-0.0163
				(0.0303)	(0.0303)

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a]Sandy soil is the comparison category; [b] Flat slope is the comparison category

Table 2.9 (cont'd)

Dep var: Soil fertility perception (1=fertile)	(1) *Individual soil properties*	(2) *Maize yield*	(3) *Plot characteristics*	(4) *Plot manager characteristics*	(5) *Area infrastructure*
Farming experience (yrs)				0.000439	0.000391
				(0.000954)	(0.000956)
Plot owned (1=yes)				-0.0563	-0.0614
				(0.0487)	(0.0486)
Walking time to plot (mins)				0.00240	0.00228
				(0.00159)	(0.00159)
Distance to extension (km)					-0.00417
					(0.00277)
Distance to town (km)					-0.00120
					(0.00112)
Distance to paved road (km)					-0.000621
					(0.00296)
Regional dummies	Yes	Yes	Yes	Yes	Yes
Observations	1268	1268	1268	1268	1268
LR χ^2	20.86	31.67	36.78	42.14	45.94
Prob > χ^2	0.0525	0.0027	0.0022	0.0040	0.0045
Pseudo R^2	0.0121	0.0184	0.0214	0.0245	0.0267

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: [a]Sandy soil is the comparison category; [b] Flat slope is the comparison category

The data shows that 99% of the sample plots had not had their soils tested during the three-year period prior to the 2014 survey when soil sampling and testing was conducted as part of the survey. This confirms that farmers' perceptions about soil fertility on their plots is largely not based on knowledge about physical, chemical and biological attributes of soil. Farmers were asked in the 2016 survey to indicate what they consider to conclude that the fertility of soil is wanting. Table 2.10 displays the frequency distribution of farmer-reported indicators of soil infertility. The statistics indicate that farmers overwhelmingly rely on crop performance, including crop health (by look), growth and yield, to judge the fertility status of the soil. Farmers also use local classifying attributes of soil such as color, texture and depth to assess soil fertility. Soil testing as a way of knowing soil fertility status comes last in the list. This can be problematic because poor crop performance may be a result of factors other than soil infertility, such as poor-quality seeds, pests, and diseases that may not be visible to the farmer. In addition, even if poor crop performance was because of soil infertility, it may not identify the exact deficiency in the soil that is contributing to infertility. Meaningful diagnosis of deficiencies in the soil can be only through soil testing. Therefore, the result that soil testing is generally a rare practice among maize farmers should raise concern about the appropriateness of the soil fertility management practices they apply. This is particularly important because the soils have been cultivated for a long time with intensity of cultivation increasing (i.e. less fallowing as inferred from the number of years of cultivation during the last decade). Indeed, it is not surprising that farmers' perceived and measured soil fertility largely mismatch in our analysis, and maize yield has stagnated or declined in the long term despite evidence of increased use of inorganic fertilizers (mainly DAP) and improved seed varieties.

Table 2.10: Farmers' reported indicators of soil infertility

Indicators	Frequency	Percent of responses	Percent of cases
unhealthy-looking plants	402	27.61	64.42
poor crop yield	399	27.40	63.94
retarded plant growth	251	17.24	40.22
soil color is different from what I expect	135	9.27	21.63
invasion by particular weeds	103	7.07	16.51
soil hard to work (hard to till)	37	2.54	5.93
texture is too coarse	34	2.34	5.45
texture is too fine	31	2.13	4.97
invasion by particular animals/organisms	27	1.85	4.33
soil is shallow	14	0.96	2.24
poor water infiltration	13	0.89	2.08
poor drainage	8	0.55	1.28
through soil testing	2	0.14	0.32
Total	1456	100	233.33

2.3.2 Adoption of soil fertility management practices

Before discussing the multivariate probit results of determinants of adoption of soil fertility management practices in Table 2.12, we discuss the correlation matrix of the equations in Table 2.11. The positive and negative signs of the correlation coefficients indicate complementarity and substitutability, respectively, in the decision to adopt the soil fertility management practices in question. The statistical significance of some of the correlation coefficients indicates simultaneity of adoption decisions and validates the suitability of multivariate probit to model adoption of the soil fertility management practices. Inorganic fertilizer and manure/compost are significantly substitutes, a finding that is consistent with Wainaina et al (2016) and Koppmair et al (2017) but differs from Marenya & Barrett (2007) and Arslan et al (2017). This suggests that an average farmer in our study may understand manure/compost and inorganic fertilizers as serving the same purpose concerning soil fertility, which is not entirely the case. In addition to supplying nutrients – mainly nitrogen – manure/compost adds organic materials into the soil and thus contributes to enhancing soil quality characteristics that are associated with increased soil organic matter, and therefore soil organic carbon, as explained earlier in the introduction. Inorganic fertilizer, on the other hand, do not add organic matter into the soil, implying that efficiency of inorganic fertilizer can be enhanced with concomitant application of manure/compost. Soil erosion control through investment in structures such as grass strips, cut-off drains and terraces are substitutes with both crops residue and inorganic fertilizer. The negative correlation of adoption between soil erosion and crops residue could be driven by mulching, which, apart from using crops vegetative remains to conserve moisture and enhance organic matter in the soil, also protects topsoil from erosion by direct impact of rain. We also see a positive correlation between adoption of manure/compost and legume intercrop.

Table 2.11: Correlation matrix for the regression equations

	Manure/compost	Crops residue	Legume intercrop	Soil erosion control
Crops residue	0.0666			
	(0.0517)			
Legume intercrop	0.102^{*}	-0.0743		
	(0.0594)	(0.0540)		
Soil erosion control	-0.0168	-0.181^{***}	0.0481	
	(0.0508)	(0.0475)	(0.0537)	
Inorganic fertilizer	-0.289^{***}	-0.0196	0.0369	-0.156^{**}
	(0.0895)	(0.0763)	(0.0829)	(0.0767)

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The complementarities and substitutabilities of the soil fertility management practices have important implications for policy and strategies to promote their use by farmers, since targeting one practice for adoption can enhance or inhibit adoption of another. In Kenya, for example, improving agricultural productivity through promotion of widespread use of inorganic fertilizers has been a major strategy by the government, evident in fertilizer subsidy programs and, lately, building of a fertilizer manufacturing plant. These initiatives, however, are not accompanied by efforts to educate farmers about and promote use of organic soil amendments, such as use of manure, compost and crops residue, which are necessary for building soil fertility and enhancing efficiency of inorganic fertilizer use. Thus, the strategy may be portraying inorganic fertilizers as the solution to the problem of soil infertility, making farmers to overemphasize use of inorganic fertilizers and undermine organic soil amendments instead of looking at both as necessary complements to enhance soil fertility for sustainable productivity improvement.

The effects of soil fertility perception on adoption shows that on average, the likelihood of applying inorganic fertilizer on a plot is significantly lower when a farmer perceives the plot to be fertile (Table 2.12). This may imply that plots that are viewed to be fertile are likely to be performing well in maize production, the farmers' main yardstick for soil fertility as we saw earlier, without application of inorganic fertilizer. The finding may also imply that on average, farmers are inclined to view inorganic fertilizer as a solution to soil infertility. This may be problematic if addressing the soil infertility problem also requires, and it does on the maize plots in this study, soil amendments that are different from what inorganic fertilizers do, such as correcting soil pH to address acidity problems and increasing organic matter to increase soil organic carbon. This is more likely what is happening in the major maize growing areas of Kenya where farmers have sustained use of inorganic fertilizers on soils that are generally low in pH and have low organic carbon levels in efforts to raise maize productivity. Soil fertility perception also has a negative but statistically weak association with adoption of maize-legume intercropping.

Among plot characteristics, slope has an increasing effect on the likelihood of soil erosion control, with erosion control structures more likely on moderate to steep relative to flat plots. Fertilizer adoption is also more likely on plots that are steep relative to those that are flat. As would be expected, plot size has a reducing effect on the probability of manure/compost use and use of crops residue. This can be attributed to the bulkiness of the practices which makes them much labor demanding and, therefore, much costlier to apply on large plots. In addition, availability of manure/compost and crops residue, which competes with other uses such as livestock feeding, to apply on large plots may be a challenge. Plot size, on the other hand, has a positive and significant association with inorganic fertilizer use.

Plot distance from the homestead is negatively correlated with adoption of manure/compost and crops residue, similar to findings of Kassie et al (2015). This may be because of the bulkiness of manure/compost, which are often made in the homestead and transported to the farm. Similar to findings by Kamau et al. (2014), Kassie et al. (2015) and Wainaina et al. (2016), plot ownership, which implies tenure security, has a positive association with adoption of manure/compost. This is expected since returns to manure/compost application is often not immediate and thus farmers are more likely to apply manure/compost on plots for which they have assurance of continued use.

Table 2.12: Multivariate probit estimation results of factors affecting adoption of soil management practices (N=1268)

	Manure/ compost		Crop residue		Legume intercrop	
	Coef.	SE	Coef.	SE	Coef.	SE
Soil fertility perception (1=fertile)	0.124	(0.0809)	-0.0286	(0.0753)	-0.159*	(0.0858)
Years of plot cultivation	0.0593*	(0.0307)	-0.0407	(0.0250)	-0.0475	(0.0290)
Plot slope (1=moderate)[a]	0.0927	(0.0900)	-0.0639	(0.0843)	0.0275	(0.0956)
Plot slope (1=steep)[a]	0.208	(0.162)	-0.201	(0.151)	0.133	(0.170)
Plot size (acres)	-0.268***	(0.0568)	-0.0546*	(0.0332)	-0.0626	(0.0403)
Walking time to plot (mins)	-0.0262***	(0.00618)	-0.0118**	(0.00467)	-0.00142	(0.00409)
Plot owned (1=yes)	0.730***	(0.154)	0.0840	(0.133)	0.139	(0.146)
Plot manager education (yrs)	0.0161	(0.0122)	-0.0249**	(0.0112)	-0.0256**	(0.0128)
Plot manager is female (1=yes)	0.0805	(0.0878)	-0.0907	(0.0825)	0.264***	(0.0957)
Farming experience (yrs)	-0.00100	(0.00286)	0.000566	(0.00266)	0.00410	(0.00305)
Active adults (#)	-0.00597	(0.0255)	0.0567**	(0.0235)	-0.0202	(0.0259)
Group membership (1=yes)	-0.0791	(0.0937)	-0.0183	(0.0881)	0.00246	(0.0998)
HH has non-farm income (1=yes)	0.0421	(0.0973)	0.122	(0.0894)	0.0913	(0.0993)
Livestock value (KES) (log)	0.105***	(0.0147)	-0.0139	(0.0132)	0.0146	(0.0145)
HH landholding (acres)	-0.000468	(0.00242)	-0.0115	(0.00871)	-0.0447***	(0.0127)
Asset index	0.636	(1.526)	2.187**	(0.868)	0.469	(1.041)
Distance to extension (km)	-0.00937	(0.00850)	-0.0153*	(0.00784)	-0.00175	(0.00878)
Distance to paved road (km)	0.00177	(0.00883)	0.00729	(0.00772)	-0.0159*	(0.00838)
Fertilizer price (KES/kg)	-0.000270	(0.00262)	-0.00391	(0.00248)	-0.00194	(0.00275)
Trans Nzoia	-0.0650	(0.172)	0.343**	(0.152)	1.157***	(0.163)
Kakamega	0.533***	(0.141)	0.531***	(0.132)	1.632***	(0.155)
Kisii	0.522***	(0.160)	1.014***	(0.150)	1.285***	(0.165)
Machakos	1.397***	(0.169)	0.120	(0.151)	1.021***	(0.161)
Constant	-2.736***	(0.433)	0.249	(0.367)	0.320	(0.411)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Log likelihood = -3137.4466; Wald chi2(115) = 1037.21; [a] Base category: flat

Table 2.12 (cont'd)

	Soil erosion control		Inorganic fertilizer	
	Coef.	SE	Coef.	SE
Soil fertility perception (1=fertile)	-0.0517	(0.0740)	-0.253**	(0.128)
Years of plot cultivation	0.0220	(0.0248)	0.0379	(0.0384)
Plot slope (1=moderate)[a]	0.447***	(0.0827)	0.165	(0.140)
Plot slope (1=steep)[a]	0.297**	(0.147)	0.566**	(0.243)
Plot size (acres)	0.0397	(0.0329)	0.295***	(0.115)
Walking time to plot (mins)	-0.000856	(0.00377)	0.0114	(0.00881)
Plot owned (1=yes)	-0.156	(0.129)	0.246	(0.198)
Plot manager education (yrs)	0.00565	(0.0110)	0.0302	(0.0190)
Plot manager is female (1=yes)	0.0321	(0.0807)	0.179	(0.137)
Farming experience (yrs)	0.00123	(0.00262)	-0.00327	(0.00401)
Active adults (#)	-0.0485**	(0.0231)	0.0502	(0.0429)
Group membership (1=yes)	-0.0326	(0.0861)	0.520***	(0.145)
HH has non-farm income (1=yes)	0.237***	(0.0869)	0.0481	(0.144)
Livestock value (KES) (log)	0.0110	(0.0129)	0.00818	(0.0185)
HH landholding (acres)	0.0163*	(0.00910)	-0.0294*	(0.0157)
Asset index	-3.241***	(1.109)	5.051	(3.504)
Distance to extension (km)	-0.00532	(0.00728)	0.0204	(0.0151)
Distance to paved road (km)	0.00178	(0.00765)	-0.0144	(0.0130)
Fertilizer price (KES/kg)	-0.00467*	(0.00242)	-0.00266	(0.00417)
Trans Nzoia	0.00887	(0.149)	0.0414	(0.254)
Kakamega	-0.0445	(0.130)	0.254	(0.233)
Kisii	0.137	(0.144)	0.978***	(0.350)
Machakos	0.489***	(0.148)	-0.672***	(0.246)
Constant	-0.334	(0.360)	0.121	(0.570)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Log likelihood = -3137.4466; Wald chi2(115) = 1037.21; [a] Base category: flat

Contrary to expectations, plot manager education has reducing effects on the likelihood of adoption of crops residue and maize-legumes intercrop and no significant effect on adoption of the other practices. Kassie et al. (2015) and Wainaina et al. (2016) found a positive effect of education of household head on adoption of manure and inorganic fertilizer while Arslan et al. (2017) found a positive effect on adoption of soil and water conservation measures. Concerning gender, female plot managers are more likely to intercrop maize with legumes compared to their male counterparts.

As expected, number of adults, a measure of labor availability, is positively correlated with adoption of crops residue, which may reflect the labor-intensive nature of the practice. However, number of adults is inversely related with soil erosion control. As expected, and similar to findings of Kassie et al. (2015), group membership, which is an indicator of social capital and may also facilitate access to information and credit, has a positive effect on inorganic fertilizer adoption, while the presence of non-farm income in a household has a positive effect on soil erosion control. As expected, livestock wealth has a positive effect on manure/compost adoption, consistent with Kassie et al. (2015) and Wainaina et al. (2016). Household landholding has an inverse relationship with maize-legume intercropping and inorganic fertilizer adoption, suggesting that smaller landholdings motivate agricultural intensification practices. Landholding has a positive association with soil erosion control while the effect of asset index is positive on use of crops residue but negative on soil erosion control.

The effects of distance to extension service provider, distance to paved road and fertilizer price are negative and weakly significant on the adoption of crops residue, maize-legume intercrop and soil erosion control, respectively. While the sign is as expected, the effect of fertilizer price on fertilizer adoption is insignificant possibly because of little variation in fertilizer prices and the high rate of adoption.

2.4 Conclusion and Implications

Soil infertility is a major problem in sub-Saharan Africa and is one of the main causes of persistently low agricultural productivity. Soil degradation because of land management practices that do not replenish soil nutrients and organic matter is the main cause of soil infertility. In Kenya, low agricultural productivity in general and poor maize yield in particular, despite notable increase in use over time of mineral fertilizers and improved maize varieties, is largely attributed to this problem. A particular problem is that continued use of chemical fertilizers, especially acidifying ones such as DAP, to soils with low pH and low organic matter can further degrade the soil. Replenishing such soils requires applying organic materials such as compost or manure and ameliorating soil acidity (Lal, 2006; Chivenge et al., 2011; Kunhikrishnan et al., 2016). Promoting recycling of organic matter into the soil and judicious use of appropriate chemical fertilizers is a necessary priority to improve agricultural productivity in general and maize yield in particular.

Several studies have explored factors that promote or hinder adoption of sustainable management practices in Kenya and elsewhere. However, often not included in many of the studies is the role of farmers' perceptions about the fertility status of their soil in the adoption choice decisions of alternative soil fertility management practices. Using household, farmer and plot-level data from major maize growing areas in Kenya, this study has estimated correspondence between maize farmers' perceived and measured soil fertility and investigated the association between farmers' perception about soil fertility and adoption (use) of soil fertility management practices.

Five key results have emerged. First, descriptive analysis results have shown that farmers' perceived and measured fertility status of 46% of the plots were consistent, about half were rated as fertile by farmers but measured to be infertile, and less than 5% were perceived as infertile by farmers but measured to be fertile. In addition, farmers overwhelmingly relied on crops performance to judge the fertility status of soil. Soil testing as a mechanism to understand the fertility status of soil was extremely rare. Secondly, interrater agreement analysis has shown little correspondence between farmers' perceived and measured soil fertility. About half of the plots were rated as fertile by farmers but measured to be infertile, suggesting substantial underestimation of soil infertility problem among the maize farmers. Thirdly, probit model results have shown that measured soil fertility and soil organic carbon and pH are significant predictors of farmers' perception about soil fertility. However, these effects vanish upon introducing in the model the maize yield variable for which the significance of the effect becomes persistent. Fourthly, multivariate probit estimation results have shown that farmers' soil fertility perception has a strong relationship with the decision to apply inorganic fertilizer, but not soil fertility management practices such as applying manure or compost. The likelihood of applying inorganic fertilizer increases when a farmer perceives the soil to be infertile. In addition, farmers appear to regard manure/compost and inorganic fertilizers as serving the same purpose concerning soil fertility, thus treating them as substitutes.

These results have important implications for policy and extension efforts to explore ways to improve management of soils for improved agricultural productivity. First, the lack of soil testing, farmers' reliance on crop performance as the main yardstick for judging soil fertility condition, and the low correspondence between farmers' perceived and measured soil fertility collectively imply that farmers are applying management practices that may not match the fertility needs of soil on their plots. This is exemplified by the result that there is persistent application of an acidifying fertilizer (DAP) to maize and low application of organic soil amendments such as manure and compost even on plots with soils that are low to very low in soil pH and organic carbon. Data has shown that over 40% of the plots that were low in soil pH and soil organic matter had DAP applied without manure or compost application.

Secondly, farmers' higher adoption rate of fertilizer relative to other organic soil amendments, their treatment of manure/compost and inorganic fertilizers as substitutes rather than complements and the strong inverse relationship between farmers' perception of soil fertility and the decision to apply inorganic fertilizer may indicate deficiencies in policy and extension information. The government promotes use of inorganic fertilizers, but without accompanying concerted efforts to educate farmers about and promote use of organic soil amendments, which are necessary for building soil fertility and enhancing efficiency of inorganic fertilizer use. This may erroneously valorize inorganic fertilizers as the solution to the problem of soil infertility and lead farmers to put less effort in the necessary complementary organic soil amendments, recognizing that manure, compost and crops residue are not readily available and are not costless to obtain and their use is labor intensive. In addition, the government continues to subsidize DAP, one of the acidifying fertilizers, for maize production, despite its own report showing that low soil pH and organic carbon levels are widespread on farms in the major maize growing areas. This policy disconnect, together with lack of aggressive extension efforts to promote organic soil amendments, may be perpetuating improper practices for soil fertility management thereby exacerbating soil degradation.

There is a need for a shift in policy and extension focus to encourage farmers to pay greater attention to organic soil amendments and judicious application of appropriate chemical fertilizers to restore and sustain soil fertility. It is acknowledged that organic soil amendments are scarce and several factors, such as ability to keep livestock for manure and labor availability for use of crops residue, determine farmers' access to and their application of them. However, there still needs to be active policy and extension efforts, perhaps like those for encouraging chemical fertilizer use, to encourage farmers to embrace their importance. In addition, policy actions that encourage judicious use of appropriate chemical fertilizers are vital. In this regard, encouraging and supporting farmers to test their soils and facilitating availability of fertilizers that are suitable for the varied soil conditions should be desirable.

REFERENCES

REFERENCES

Ali, A., & Abdulai, A. (2010). The Adoption of Genetically Modified Cotton and Poverty Reduction in Pakistan. *Journal of Agricultural Economics*, *61*(1), 175–192.

Ariga, J., & Jayne, T. S. (2009). *Private Sector Responses to Public Investments and Policy Reforms The Case of Fertilizer and Maize Market Development in Kenya* (IFPRI Discussion Paper No. 921).

Arslan, A., Belotti, F., & Lipper, L. (2017). Smallholder productivity and weather shocks: Adoption and impact of widely promoted agricultural practices in Tanzania. *Food Policy*, *69*, 68–81. https://doi.org/10.1016/j.foodpol.2017.03.005

Asafu-Adjaye, J. (2008). Factors affecting the adoption of soil conservation measures: A case study of Fijian cane farmers. *Journal of Agricultural and Resource Economics*, *33*(1), 99–117.

Berazneva, J., Mcbride, L., Sheahan, M., & David, G. (2016). Perceived , measured , and estimated soil fertility in east Africa : Implications for farmers and researchers. In *2016 Annual Meeting, July 31-August 2, 2016*. Boston, Massachusetts: Agricultural and Applied Economics Association.

Chivenge, P., Vanlauwe, B., & Six, J. (2011). Does the combined application of organic and mineral nutrient sources influence maize productivity ? A meta-analysis. *Plant Soil*, *342*, 1–30.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *XX*(1), 37–46.

Corbeels, M., Shiferaw, A., & Haile, M. (2000). *Farmers ' knowledge of soil fertility and local management strategies in* (Managing Africa's Soils No. 10).

de Janvry, A., Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets : Some Paradoxes Explained. *The Economic Journal*, *101*(409), 1400–1417.

Desbiez, A., Matthews, R., Tripathi, B., & Ellis-Jones, J. (2004). Perceptions and assessment of soil fertility by farmers in the mid-hills of Nepal. *Agriculture, Ecosystems and Environment*, *103*(1), 191–206.

Doran, J. W., & Parkin, T. B. (1994). Defining and assessing soil quality. *Defining Soil Quality for a Sustainable Environment*.

Eckert, E., & Bell, A. (2005). Invisible Force : Farmers â€™ Mental Models and How They Influence Learning and Actions. *Journal of Extension*, *43*(3), 1–7.

Enyong, L. A., Debrah, S. K., & Bationo, A. (1999). Farmers ' perceptions and attitudes towards introduced soil-fertility enhancing technologies in western Africa. *Nutrient Cycling in Agroecosystems*, *53*, 177–187.

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of Agricultural Innovations in Developing Countries : A Survey. *Economic Development and Cultural Change*, *33*(2), 255–298.

Fleiss, J. L., Levine, B., & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc.

Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, *2*(1), 395–424.

Government of Kenya. (2014). *Soil suitability evaluation for maize production in Kenya soil suitability evaluation for*. Nairobi, Kenya.

Grisso, R., Alley, M., Holshouser, D., & Thomason, W. (2009). Precision farming tools: soil electrical conductivity. *Virginia Cooperative Extension*, *442*(508), 1–6. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Precision+farming+tools:+soil+electrical+conductivity#7

Halbrendt, J., Gray, S. A., Crow, S., Radovich, T., Kimura, A. H., & Tamang, B. B. (2014). Differences in farmer and expert beliefs and the perceived impacts of conservation agriculture. *Global Environmental Change*, *28*(1), 50–62.

Horneck, D. A., Sullivan, D. M., Owen, J. S., & Hart, J. M. (2011). *Soil Test Interpretation Guide* (EC No. 1478). Corvallis, OR.

IIASA/FAO. (2012a). *Global Agro-ecological Zones (GAEZ v3.0)*. Laxenburg, Austria and Rome, Italy.

IIASA/FAO. (2012b). *Global Agro-ecological Zones (GAEZ v3.0)*. *IIASA, Laxenburg, Austria and FAO, Rome, Italy. iv*. https://doi.org/10.1029/97GB03657

Jack, B. K. (2011). *Market inefficiencies and the adoption of agricultural technologies in developing countries*. *Literature review, Agricultural Technology Adoption Initiative*.

Kamau, M., Smale, M., & Mutua, M. (2014). Farmer demand for soil fertility management practices in Kenya ' s grain basket. *Food Security*, 793–806.

Kassie, M., Shiferaw, B., & Muricho, G. (2011). Agricultural Technology , Crop Income , and Poverty Alleviation in Uganda. *World Development*, *39*(10), 1784–1795.

Kassie, M., Teklewold, H., Jaleta, M., Marenya, P., & Erenstein, O. (2015a). Understanding the adoption of a portfolio of sustainable intensification practices in eastern and southern Africa. *Land Use Policy*, *42*, 400–411. https://doi.org/10.1016/J.LANDUSEPOL.2014.08.016

Kassie, M., Teklewold, H., Jaleta, M., Marenya, P., & Erenstein, O. (2015b). Understanding the adoption of a portfolio of sustainable intensification practices in eastern and southern Africa.

Land Use Policy, *42*. https://doi.org/10.1016/j.landusepol.2014.08.016

Kelly, V., Adesina, A. A., & Gordon, A. (2003). Expanding access to agricultural inputs in Africa: A review of recent market development experience. *Food Policy*, *28*(4), 379–404.

Kerr, J., & Pender, J. (2005). Farmers ' perceptions of soil erosion and its consequences in India ' s semiarid tropics. *Land Degradation & Development*, *16*, 257–271.

Knowler, D., & Bradshaw, B. (2007). Farmers ' adoption of conservation agriculture : A review and synthesis of recent research. *Food Policy*, *32*, 25–48.

Koppmair, S., Kassie, M., & Qaim, M. (2017). The influence of farm input subsidies on the adoption of natural resource management technologies. *Australian Journal of Agricultural and Resource Economics*, *61*, 539–556. https://doi.org/10.1111/1467-8489.12220

Kunhikrishnan, A., Thangarajan, R., Bolan, N. S., Xu, Y., Mandal, S., Gleeson, D. B., … Naidu, R. (2016). Functional Relationships of Soil Acidification, Liming, and Greenhouse Gas Flux. *Advances in Agronomy*, *139*, 1–71. https://doi.org/10.1016/bs.agron.2016.05.001

Lal, R. (2006). Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, *209*(August 2005), 197–209.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Marenya, P., Barrett, C. B., & Gulick, T. (2008). Farmers' perceptions of soil fertility and fertilizer yield response in Kenya. Retrieved from http://ssrn.com/abstract=1845546

Marenya, P. P., & Barrett, C. B. (2007). Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. *Food Policy*, *32*(4), 515–536.

Marenya, P. P., & Barrett, C. B. (2009). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, *40*, 561–572.

McCauley, A., Jones, C., & Olson-Rutz, K. (2017). Soil pH and organic matter. *Nutrient Management*, (8), 1–16.

Odendo, M., Obare, G., & Salasya, B. (2009). Factors responsible for differences in uptake of integrated soil fertility management practices amongst smallholders in western Kenya. *African Journal of Agricultural Research*, *4*(11), 1303–1311.

Okalebo, J. R., Gathua, K. W., & Woomer, P. L. (2002). *Laboratory methods of soil and plant analysis: A working manual* (Technical Publication No. Volume 1).

Pender, J. L., & Kerr, J. M. (1998). Determinants of farmers' indigenous soil and water conservation investments in semi-arid India. *Agricultural Economics*, *19*.

Sanchez, P. A., Shepherd, K. D., Soule, M. J., Place, F. M., Buresh, R. J., Izac, A. N., & Kwesiga, F. R. (1997). Soil Fertility Replenishment in Africa : An Investment in Natural Resource Capital. *American Society of Agronomy and Soil Science Society of America*, (51), 1–46.

Singh, I., Squire, L., & Strauss, J. (1986). A Survey of Agricultural Household Models : Recent Findings and Policy Implications. *The World Bank Economic Review*, *1*(1), 149–179.

Smale, M., & Olwande, J. (2014). Demand for maize hybrids and hybrid change on smallholder farms in Kenya. *Agricultural Economics (United Kingdom)*, *45*(4).

Smaling, E. M. A., Nandwa, S. M., & Janssen, B. H. (1997). Soil Fertility in Africa Is at Stake. *American Society of Agronomy and Soil Science Society of America*, (51).

Stoorvogel, J. J., & Windmeijer, P. N. (1993). Calculating soil nutrient balances in Africa at different scales. *Fertilizer Research*, *35*, 237–250.

Stoorvogel, J., & Smaling, E. M. A. (1998). Research on soil fertility decline in tropical environments : integration of spatial scales. *Nutrient Cycling in Agroecosystems 50:*, *50*, 151–158.

The Montepellier Panel. (2013). *Sustainable Intensification: A New Paradigm for African Agriculture*. London.

Tittonell, P., Corbeels, M., Wijk, M. T. Van, Vanlauwe, B., & Giller, K. E. (2008). Combining Organic and Mineral Fertilizers for Integrated Soil Fertility Management in Smallholder Farming Systems of Kenya: Explorations Using the Crop-Soil Model FIELD. *Agronomy Journal*, *100*(5), 1511–1526.

Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, *143*, 76–90. https://doi.org/10.1016/j.fcr.2012.10.007

Turner, B. L., Wuellner, M., Nichols, T., & Gates, R. (2014). Dueling Land Ethics : Uncovering Agricultural Stakeholder Mental Models to Better Understand Recent Land Use Conversion. *Journal of Environmental Ethics*, (27), 831–856.

Udry, C. (2010). The economics of agriculture in Africa: Notes toward a research program. *AfJARE*, *5*(1), 284–299.

Vignola, R., Koellner, T., Scholz, R. W., & McDaniels, T. L. (2010). Decision-making by farmers regarding ecosystem services: Factors affecting soil conservation efforts in Costa Rica. *Land Use Policy*, *27*(4), 1132–1142.

Wainaina, P., Tongruksawattana, S., & Qaim, M. (2016). Tradeoffs and complementarities in the adoption of improved seeds , fertilizer , and natural resource management technologies in Kenya. *Agricultural Economics*, *47*, 351–362.

Willy, D. K., Muyanga, M., & Jayne, T. (2016). Can economic and environmental benefits associated with agricultural intensification be sustained at high population densities ? A farm level empirical analysis. In *Invited paper presented at the 5th International Conference of the African Association of Agricultural Economists, September 23-26, 2016, Addis Ababa, Ethiopia*.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data. *MIT Press*. https://doi.org/10.1515/humr.2003.021

CHAPTER 3: TECHNICAL EFFICIENCY AND SOIL FERTILITY AND AGRONOMIC PRACTICES IN MAIZE PRODUCTION IN KENYA

3.1 Introduction and Study Objectives

There has been a renewed interest about the role of agriculture in economic development in recent years, with increased recognition that it is key to sustainable development and poverty reduction in agriculture-based economies. This is particularly true of developing countries in sub-Saharan Africa where over three quarters of the population is rural and majority of who depend almost entirely on agriculture for livelihood (The World Bank, 2007; Haggblade & Hazell 2010). Sustainable agricultural productivity growth is thus essential to broad-based sustainable economic progress in these countries (The World Bank, 2007; Diao et al (2007), and fostering that growth requires technological progress and/or efficiency improvement in management of productive resources.

The problem of low agricultural productivity in sub-Saharan Africa has been majorly attributed to soil infertility due to nutrient depletion (Stoorvogel & Smaling, 1998; Sanchez et al., 1997; The Montepellier Panel, 2013; Smaling et al., 1997; Stoorvogel & Windmeijer, 1993). Soil nutrient depletion is blamed on unsustainable land management practices that deplete soil organic carbon pool (Lal, 2006) and accelerate soil acidification (Kunhikrishnan et al., 2016). Soils that are in such depleted conditions result in reduced response of crops to external inputs such as fertilizers, irrigation and high yielding crop varieties (Lal, 2006; Tittonell & Giller, 2013). This in turn lowers productivity and reduces efficiency of input use.

In Kenya, soil degradation in terms of nutrient imbalance, diminished soil organic carbon pool and high acidity is one of the main reasons cited for lack of growth in yield of maize (Government of Kenya, 2014; Tittonell et al, 2008; Marenya & Barrett, 2007). Farmers have intensified use of external inputs such as chemical fertilizers and high yielding maize varieties over time (e.g. Ariga & Jayne, 2009, 2009; Smale & Olwande, 2014; Muyanga & Jayne, 2014), but evidence suggests that aggregate maize yield has generally stagnated or declined. This implies that soil infertility combined with agronomic practices that may not effectively respond to fertility needs of the soil could be affecting maize response to external inputs.

Because there is little scope for land expansion in Kenya, the onus of increasing agricultural production, particularly maize, is on increasing productivity, i.e. output per given amounts of inputs. This can be achieved through technological change, increased efficiency in use of existing technology and productive resources or both. This study is concerned with efficiency for two reasons. First, when soils are degraded and farmers do not apply proper agronomic practices in response to the actual conditions of the soil, they are unlikely to realize maximum yield possible with the amount of inputs used under a given production technology. This means they are likely to be technically inefficient. In this respect, knowledge of the level of technical efficiency of maize farmers would be useful in determining potential productivity gains possible through better agronomic management under existing technology. Information about specific factors that affect variations in efficiency across farms can guide targeted policies to improve efficiency for wide-scale maize productivity improvement.

Secondly, higher yields because of increased amount of inputs may not necessarily imply higher technical efficiency. A farmer may apply larger amounts of inputs and achieve higher yields but remain far below the maximum yield possible with the amount of the inputs applied (Mochebelele

& Winter-Nelson, 2000). Therefore, improving technical efficiency is important for sustainable agricultural intensification not only economically but also environmentally, since improper use of certain external inputs such as inorganic fertilizers, pesticides and herbicides may potentially degrade the environment and its productive resources.

Using a stochastic production frontier (SF) approach due to Meeusen & van Den Broeck (1977) and Aigner et al (1977), this study estimates technical efficiency of maize farmers in Kenya while controlling for soil fertility conditions and agronomic practices, and identifies factors responsible for heterogeneity in technical efficiency across farms. The study addresses the following specific questions:

(a) What is the technical efficiency level of maize farmers? How does technical efficiency vary across farms? What does this mean for measures to improve maize productivity?

(b) What factors explain technical efficiency variations across farms?

(c) How important are soil fertility conditions and agronomic practices to agricultural productivity and efficiency estimation?

A major contribution of this study is that it evaluates maize productivity and efficiency while integrating soil fertility conditions and agronomic practices in the estimation, thus overcoming omitted variable bias to which many studies that do not include these conditions and practices may be prone. While studies have been conducted in Kenya and elsewhere on agricultural productivity and efficiency, such studies have often inadequately specified or ignored environmental variables mainly because of data limitations (e.g. Liu & Myers, 2009 and Kibaara, 2005 in Kenya; Mochebelele & Winter-Nelson, 2000; Tiedemann & Latacz-Lohmann, 2013). Ali & Byerlee

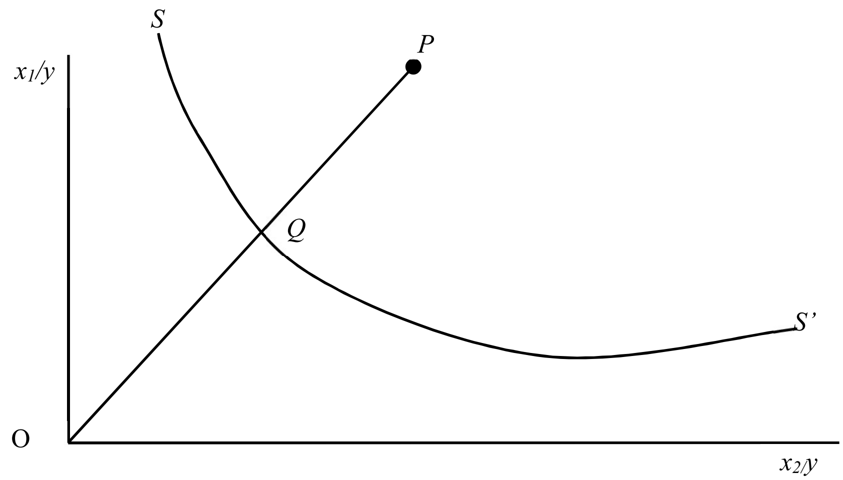
(1991) suggest and Sherlund et al (2002) and Ekbom et al (2013) have shown that omitting environmental production conditions from farm technical efficiency analysis would result in biased and inconsistent estimates of the production frontier parameters and understated technical efficiency. Maize-growing population in Kenya is quite heterogeneous and diversified, with biophysical conditions under which farmers operate among the most important sources of heterogeneity and diversity. These conditions affect farmers' input choice decisions hence the need to control for them in productivity and efficiency analysis.

3.3 Methods and Data

3.3.1 Conceptual Framework

The conceptual framework for technical efficiency is based on the seminal work of Farrell (1957). In the context of a firm, Farrell (1957) defined technical efficiency (TE) as a firm's ability to maximize its production from a given level of inputs under a given technology. Farrell (1957) illustrated this concept as in Figure 1 below, obtained from Coelli et al (2005). A firm is assumed to produce one output, y, using two inputs, x_1 and x_2, under constant returns to scale technology. Let *SS'* represent the isoquant of a technically efficient firm. The isoquant represents the output level that the perfectly efficient firm would produce given any combination of the two inputs, x_1 and x_2. A firm that produces at point P is technical inefficient because the input amounts can be proportionally reduced while the output level is maintained. The distance QP represents the firm's technical inefficiency. The ratio *QP/OP* measures technical inefficiency of the firm. It is the proportion by which all inputs could be downscaled to the technically efficient output level. The ratio is one for a perfectly technically efficient firm. Technical efficiency is expressed as $TE = 1 - QP/OP = OQ/OP$.

Figure 3.1: Input-oriented technical efficiency illustration (Coelli et al 2005)



This representation of the concept of technical efficiency is input-oriented and has input-reducing focus (Coelli et al (2005). A useful analogous representation of the concept is in output-oriented fashion, which views technical efficiency in terms of how much output could be proportionally increased without reducing input amounts under the same technology. Consider a single input, x, and single output, y, and decreasing returns to scale production technology, $f(x)$, in Figure 3.2 (Coelli et al, 2005). A firm producing at point P is technically inefficient since the maximum output possible with C amount of input is at D, which is on the production frontier. The ratio CP/CD measures the technical efficiency of the firm producing.

The case of two outputs, y_1 and y_2, a single input, x, and constant returns to scale technology is represented by Figure 3.3 (Coelli et al, 2005). Point A represents technically inefficient output combination because for the given level of the input it is below the maximum possible output combination, B, which is on the production possibility frontier (PPF) represented by the curve ZZ'. The ratio OA/OB measures technical efficiency of a firm producing at point A. Using the notation of Coelli et al (2005), a firm's technical efficiency can also be generally represented as the distance function at the firm's input and output vectors, $d_o(\mathbf{x}, \mathbf{y})$, a representation that is useful in the case of multiple outputs and inputs.

Figure 3.2: Output-oriented technical efficiency with one input and one output (Coelli et al 2005)

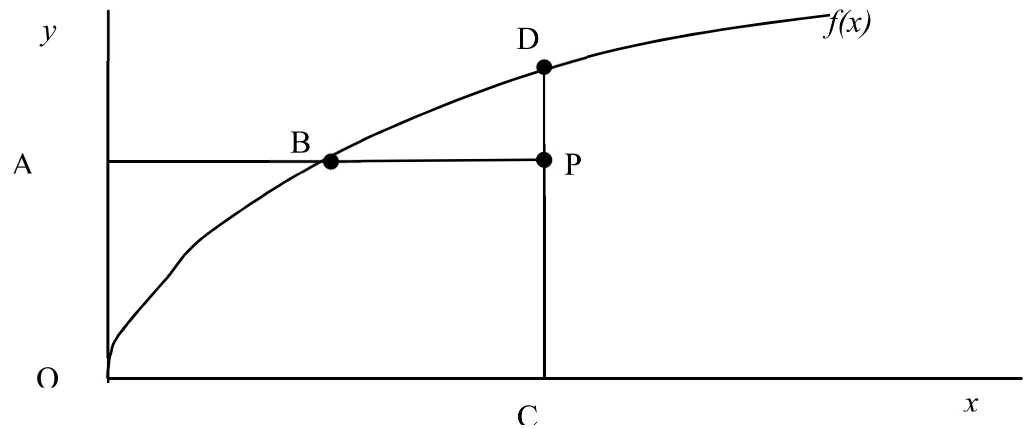
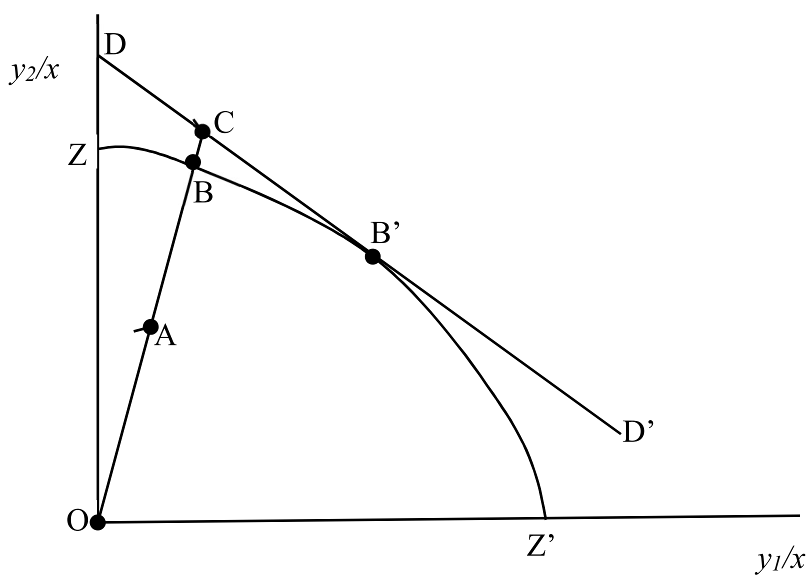


Figure 3.3: Output-oriented technical efficiency with two outputs and one input (Coelli et al 2005)



The next sub-section formalizes the framework for estimating a stochastic frontier production function in the context of a single output and multiple inputs, the relevant context for this study.

3.3.2 Stochastic frontier analysis

Production is a process that transforms a set of inputs, $\mathbf{x} \in \mathrm{R}_{\mathrm{N}}^{+}$, into a set of outputs, $\mathbf{y} \in \mathrm{R}_{\mathrm{M}}^{+}$, through some technology represented by a production function. Figure 3.2 above is a graph of production technology for the case of a single input and a single output. The production possibilities set is the input-output space bounded above by the production frontier, *f(x)*. It is the set of input-output combinations feasible with the production technology. The production possibilities frontier, which is the upper boundary of the production possibilities set, is the set of maximum output levels that can be produced from any given vector of inputs. Formally, the production frontier can be represented as:

$$f(\mathbf{x}) = \max\{y:\ y\ \in P(\mathbf{x})\} \qquad (3.1),$$

where $P(\mathbf{x})$ is the production possibilities set, i.e. the set of vectors of output that is feasible to produce for each vector of inputs.

The production frontier is the point of reference in determining technical efficiency of a producing unit as illustrated in the conceptual framework. The input-output combination of each producing unit, a maize farm in the context of this study, is either below or on the production frontier. Initially, technical efficiency analysis was concerned with measuring how far below the production frontier each producing unit operates. Later, the analysis has also often been concerned about factors that explain variations in technical efficiency across producing units, to provide information that can better guide policy interventions to improve efficiency.

There are two commonly used approaches to technical efficiency analysis. One approach is the data envelopment analysis (DEA), which is non-parametric and uses mathematical programming to estimate the efficient production frontier against which individual producing units' outputs are measured. A major criticism against DEA is that it does not separate inefficiency from random error and considers all deviation from the frontier as inefficiency (Coelli 1995). Nevertheless, DEA has the advantage that it neither requires a functional form for the production frontier nor makes any distributional assumptions (Coelli 1995). This approach has not been widely applied in efficiency studies in agriculture. The other approach is the stochastic frontier analysis (SFA), which relies on functional form specification for the production frontier and distributional assumptions about the error terms in the model (Coelli 1995). It applies econometric methods to data on producing units to estimate parameters of the production frontier and inefficiency. Unlike the DEA, the SFA approach explicitly separates inefficiency from random error.

This study applies the SFA approach to two cross sections of data on maize farms in Kenya to estimate technical efficiency and identify factors that explain variations in technical efficiency across farms. Stochastic frontier (SF) approach to technical efficiency analysis was originally developed independently by Meeusen & van Den Broeck (1977) and Aigner et al (1977). The idea and general setup of the stochastic production frontier in a cross-sectional data setting is as follows. Suppose we have a sample of maize producing farms indexed by i=1, 2,…, N and y_i and $\mathbf{x}_{\mathrm{i}}$, respectively, represent the output and a vector of inputs and other variables that affect the frontier output. Let $y_i^* \geq y_i$ be the unobserved frontier output. We can define the frontier output as:

$$y_i^* = f(\mathbf{x}_{\mathrm{i}}; \boldsymbol{\beta}) + \mathrm{v}_{\mathrm{i}}, \tag{3.2}$$

where $f(\mathbf{x}_{\mathrm{i}}; \boldsymbol{\beta})$ represents the production technology defining the frontier, $\boldsymbol{\beta}$ is the parameter vector to be estimated and v_{i} is statistical noise term which captures measurement and specification errors. The observed (actual) output is defined as the frontier output minus a non-negative error term, $u_i \geq 0$, representing the farm's inefficiency:

$$y_i = y_i^* - u_i, \tag{3.3}$$

If we express the actual and frontier output in terms of their natural logs, we can re-arrange (3.3) and represent a farm's technical efficiency measure as the ratio of actual to the frontier output as shown in (4):

$$\exp(-u_i) = \frac{y_i}{y_i^*} \tag{3.4}$$

The stochastic production frontier model is as follows:

$$y_i = f(\mathbf{x}_i; \boldsymbol{\beta}) + \varepsilon_i, \tag{3.5}$$

$$\varepsilon_i = v_i - u_i \tag{3.6}$$

$$v_i \sim N(0, \sigma_v^2), \tag{3.7}$$

$$u_i \sim \mathcal{F} \tag{3.8}$$

The error term, ε_i, has two components – the random error term, v_i, and the non-negative error term, u_i, which represents inefficiency. It is assumed that the two components of the composed error term are independent of each other and of $\mathbf{x}_i$. Aigner et al (1977) pointed out that the inefficiency, representing deviation of a firm's output from the frontier, originates from factors under the control of the firm, such as management effort. The random error, v_i, indicates that the frontier is stochastic and is a result of events that are not under a firm's control and that can be favourable and/or unfavourable, such as environmental shocks. The random error also results from measurement and observation errors.

The main objective of SFA is to disentangle inefficiency (u_i) from random error (v_i) in the composed error term. To achieve this, estimation of the model often relies on assumptions about statistical distributions of the random error, v_i, and the inefficiency term, u_i, and usually uses maximum likelihood (ML) procedure. Meeusen & van Den Broeck (1977) and Aigner et al (1977) both assumed a zero-mean normal distribution for v_i, which has been widely maintained in SFA studies. For the inefficiency component, u_i, Meeusen & van Den Broeck (1977) assumed exponential distribution with a single parameter, i.e. $u_i \sim \mathcal{E}(\sigma_u)$, while Aigner et al (1977) assumed half normal distribution, i.e. $u_i \sim N^+(0, \sigma_u^2)$. The next step in ML estimation after making distributional assumptions is to derive the log-likelihood function of the model, which is thereafter maximized with respect to the parameters and parameter values obtained.

Based on their assumption of mutual independence of v_i and u_i and their distributions, Aigner et al (1977) derived the density function of the composed error term, ε_i, and obtained the log-likelihood function of the *i*th observation as follows:

$$L_i = -ln\left(\frac{1}{2}\right) - \frac{1}{2} ln(\sigma_v^2 + \sigma_u^2) + ln\phi\left(\frac{\varepsilon_i}{\sqrt{\sigma_v^2+\sigma_u^2}}\right) + ln\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right) \tag{3.9}$$

where $\mu_{*i} = \frac{-\sigma_u^2 \varepsilon_i}{\sigma_v^2+\sigma_u^2}$; $\sigma_*^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2+\sigma_u^2}$; $\phi(\cdot)$ is a standard normal density function; and $\Phi(\cdot)$ is a standard normal cumulative distribution function (cdf). Upon maximizing the sum of the log-likelihood function over individual observations, the parameter estimates, i.e. $\left(\hat{\boldsymbol{\beta}}, \hat{\varepsilon}, \hat{\sigma}_v^2, \hat{\sigma}_u^2\right)$, and subsequently the values of $\hat{\mu}_{*i}$ and $\hat{\sigma}_*^2$ are obtained.

The model parameters are then used to derive technical inefficiency index as $E(u_i|\varepsilon_i)$ according to Jondrow et al (1982) and technical efficiency scores as $E(\exp(-u_i)|\varepsilon_i)$ according to Battese &

Coelli (1988). Jondrow et al (1982) showed the distribution of u_i conditional on ε_i to be truncated normal with mean μ_{*i} and variance σ_*^2, i.e. $(u_i|\varepsilon_i) \sim N^+(\mu_{*i}, \sigma_*^2)$. Based on this distribution, the technical ineffiency scores by Jondrow et al (1982) and efficiency scores by the Battese & Coelli (1988) methods, respectively, are predicted as:

$$\hat{u}_i = E(u_i|\varepsilon_i) = \frac{\sigma_* \phi\left(\frac{\mu_{*i}}{\sigma_*}\right)}{\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)} + \mu_{*i} \tag{3.10}$$

$$E(exp(-u_i)|\varepsilon_i) = \exp(-\mu_{*i} + \frac{1}{2}\sigma_*^2) \frac{\phi\left(\frac{\mu_{*i}}{\sigma_*} - \sigma_*\right)}{\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)} \tag{3.11}$$

There have been tremendous developments in methods and empirical applications of SFA following the initial studies by Meeusen & van Den Broeck (1977) and Aigner et al (1977). Kumbhakar & Lovell (2000 and Parmeter & Kumbhakar (2014 provide detailed accounts of the developments. Early studies focused on finding flexible distributions for the inefficiency term. For example, Stevenson (1980) proposed gamma distribution and a truncated (at 0) normal distribution, which allowed for zero and non-zero modes in the distribution of inefficiency, while Greene (1980) proposed a gamma distribution for the inefficiency term.

Jondrow et al (1982) developed the method for estimating observation-specific technical inefficiency parameters as discussed above, overcoming shortcomings in earlier methods that estimated only average inefficiency score over a sample of producing units. However, the Jondrow et al (1982) method assumes that the technical inefficiency parameter is independently and identically distributed. This assumption is problematic in cases where technical inefficiency is correlated with characteristics of producing units.

Kumbhakar, Ghosh, & McGuckin (1991), Reifschneider & Stevenson (1991), Huang & Liu (1994) and Battese & Coelli (1995) extended the stochastic frontier methodology to estimate the production frontier and explicit relationship between inefficiency and exogenous factors in a single step. This estimation procedure solves the bias problem inherent in the two-step estimation process which had dominated in earlier studies (see Wang & Schmidt (2002) for a discussion about the bias in a two-step procedure). Caudill & Ford (1993), Caudill, Ford, & Gropper (1995), Hadri (1999) and Wang (2002, 2003) considered stochastic frontier models that address heteroscedasticity problem through parameterizing the pre-truncated variance of the inefficiency term, u_i, and noted that ignoring heteroscedasticity may lead to biased inefficiency estimates. This contrasts with the case in linear models where heteroscedasticity in the error term may affect only the precision of coefficient estimates.

Wang & Schmidt (2002) introduced the idea of scaling property in the context of cross-sectional data while Alvarez et al (2006) discuss models with the property but in the context of panel data and identify its advantages. Models with the scaling property have the feature that changes in the exogenous factors hypothesized to affect inefficiency change the scale but not the distribution of the inefficiency component of the error term. Alvarez et al (2006) propose a procedure for selection of the inefficiency model.

The first studies to apply panel data methods to stochastic frontier analysis were Pitt & Lee (1981) and Schmidt & Sickles (1984), both of which assumed time-invariant technical inefficiency. The time-invariant assumption would be acceptable in short panels but obviously is practically questionable in long period panels. Another implication of this assumption is that unobserved producer-specific effects that are time-invariant are interpreted as inefficiency, which may bias estimated inefficiency scores. An additional shortcoming of the conventional fixed-effects model is that it requires sufficient variability of data over time for each producing unit because it uses within estimation. Little within-variability of data for producing units would produce imprecise fixed-effects estimates. In addition, it is not possible to include time-invariant covariates, such as gender, in the model.

To overcome some of the limitations of the conventional panel data methods, later studies using panel data relaxed the time-invariant assumption and considered time-varying inefficiency (e.g. Kumbhakar, 1990; Cornwell, Schmidt, & Sickles, 1990; Battese & Coelli, 1995). However, the models used in these studies still confound time-invariant unobserved producer-specific effects into inefficiency. In efforts to overcome this problem, Greene (2005) extended panel data methods to the stochastic frontier models and developed what he calls "true fixed-effects" (TFE) and "true random-effects" (TRE) stochastic frontier models. A key feature of these models is that they treat inefficiency as time-varying and at the same time separate time-invariant unobserved producer-specific effects from inefficiency. The TFE stochastic frontier model, however, faces the incidental parameters problem (Neyman & Scott, 1948). This problem arises because the number of fixed-effect parameters (the incidental parameters) to be estimated increases as the sample size increases while the time dimension is fixed, resulting in inconsistent estimates of the incidental parameters. Greene (2005) found that while the incidental parameter problem does not affect the model slope parameters, it biases model residuals, therefore affecting technical inefficiency estimates, which are based on the residuals. Some analysts, for example Wang & Ho (2010), Chen, Schmidt, & Wang (2014) and Belotti & Ilardi (2015), have proposed estimation methods to resolve this problem.

While its estimation is not problematic, a limitation of the TRE stochastic frontier model is that it does not allow for correlation between time-invariant unobserved producer-specific effects and explanatory variables in the model, which can lead to biased estimates. Some analysts have proposed adjustments to the TRE specification of stochastic frontier model using Mundlak-Chamberlain device (Mundlak, 1978; Chamberlain, 1984; Wooldridge, 2010) to account for correlation between time-invariant unobserved producer-specific effects and covariates in the model (e.g. Farsi, Filippini, & Kuenzle, 2005; Abdulai & Tietje 2007; Filippini, Hunt, & Zorić 2014; Filippini & Greene, 2016; Griffiths & Hajargasht, 2016). The authors note that unlike in the random effects linear model with a normal error term where adjustment using the Mundlak-Chamberlain device results in coefficient estimates that are similar to within estimators, the coefficient estimates from the adjusted TRE stochastic frontier model are not similar to within estimators. This is because the adjusted TRE stochastic frontier model has asymmetric composed error term and estimation is by maximum likelihood. Nevertheless, the authors suggest that using the device in TRE stochastic frontier model reduces heterogeneity bias because some correlation between time-invariant unobserved producer-specific effects and explanatory variables is captured. However, what remains unknown is the extent to which the heterogeneity bias is reduced.

Despite the methodological efforts in SFA to disentangle time-invariant unobserved producer-specific effects from inefficiency, Chen et al. (2014) observe that a philosophical question remains as to whether these effects are part of inefficiency, as assumed by the models that do not distinguish them from inefficiency, or heterogeneity to control for in estimating inefficiency. Some analysts have recently suggested that the 'true' inefficiency measure may lie in between the two extreme positions and proposed extensions of stochastic frontier panel data models to disentangle persistent inefficiency from time-invariant unobserved producer-specific effects (e.g. Colombi, Martini, & Vittadini, 2011; Kumbhakar, Lien, & Hardaker, 2014; Filippini & Greene, 2016).

The methodological developments in SFA are impressive, but great variability exists in results obtained from using different models. For example, Kumbhakar, Lien, & Hardaker (2014) apply six panel data models in SFA of technical efficiency in Norwegian grain farming and obtain quite different results for each of the models. They conclude that there is no model that can be said to measure inefficiency "correctly". This places the onus of model choice in an empirical study using SFA on a careful understanding of the research context and the nature of data available.

We present model choice and estimation strategy next.

3.3.3 Model choice and estimation strategy

3.3.3.1 Model choice

The focus of the study is on estimating farm-level technical efficiency while controlling for environmental conditions, and identifying factors that affect technical efficiency variation across farms. It exploits some of the methodological advances discussed above and applies them to pooled cross sectional data to meet these objectives. Reflecting on the objectives and the data, the study applies the model proposed by Wang (2002), which combines features of the model of Kumbhakar, Ghosh, & McGuckin (1991), Huang & Liu (1994) and Battese & Coelli (1995) (KGMHLBC) and that of Caudill & Ford (1993), Caudill, Ford, & Gropper(1995) and Hadri (1999) (CFCFGH). Indexing producing units (maize farms in this case) by $i = 1,2, \dots, N$, we express the model in a cross-sectional setting as follows:

$$y_i = f(\mathbf{x}_\mathrm{i}, \mathbf{w}_\mathrm{i}; \boldsymbol{\beta}, \boldsymbol{\theta}) + \varepsilon_i \tag{3.12}$$

$$\varepsilon_i = v_i - u_i, \qquad u_i \geq 0 \tag{3.13}$$

$$v_i \sim N(0, \sigma_v^2), \tag{3.14}$$

$$\mathrm{u_i} \sim N^+(\mu_i, \sigma_{ui}^2), \tag{3.15}$$

$$\mu_i = \boldsymbol{z}_i \boldsymbol{\delta}, \tag{3.16}$$

$$\sigma_{ui}^2 = \exp(\boldsymbol{z}_i \boldsymbol{\tau}), \tag{3.17}$$

where y_i is the output, $f(\mathbf{x}_\mathrm{i}, \mathbf{w}_\mathrm{i}; \boldsymbol{\beta}, \boldsymbol{\theta})$ represents the production frontier, $\mathbf{x}_\mathrm{i}$ is the vector of inputs and includes one, $\mathbf{w}_\mathrm{i}$ is a vector of environmental conditions (soil fertility conditions and agronomic practices) and $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are the corresponding parameter vectors to be estimated. The term v_i represents random error that accounts for measurement and specification errors and u_i is a non-negative random term representing inefficiency. A zero-mean normal distribution is assumed for v_i (3.14) while u_i is assumed to have a truncated normal distribution with mean μ_{it} and variance σ_{ui}^2 (3.15) both of which are functions of $\boldsymbol{z}_i$((3.16), (3.17)). The vector $\boldsymbol{z}_i$(which includes one among the elements), represents exogenous factors hypothesized to influence inefficiency, while $\boldsymbol{\delta}$ and $\boldsymbol{\tau}$ are the corresponding parameter vectors to be estimated. It is assumed that v_i and u_i are independent of each other and of $\mathbf{x}_\mathrm{i}$ and $\mathbf{w}_\mathrm{i}$.

The KGMHLBC and CFCFGH models differ in that the former allows the exogenous factors, $\boldsymbol{z}_i$, to influence inefficiency through the mean, μ_i, of the pre-truncated distribution of $u_i$while the latter accounts for heteroscedasticity through the variance of u_i, that is, the exogenous factors are allowed to influence inefficiency through the variance, σ_{ui}^2, of the pre-truncated distribution of u_i. Wang (2002) demonstrates that allowing the exogenous factors to influence inefficiency through both the mean and the variance allows for a non-monotonic relationship between inefficiency and exogenous factors. In addition, Kumbhakar & Lovell (2000) note that disregard for heteroscedasticity results in biased estimates of the production frontier parameters and technical ineffciency estimates.

It is important to acknowledge the limitation that the modelling approach used here does not accommodate the possibility that some maize farms may be fully efficient. Rho & Schmidt (2015) suggest that if some producing units are fully efficient, the traditional stochastic frontier model we apply is mis specified and may result in biased production frontier and inefficiency estimates. Although this study does not apply them, recent efforts to accommodate the possibility of some units being fully efficient include Kumbhakar, Parmeter, & Tsionas (2013), Tran & Tsionas (2015) and Rho & Schmidt (2015).

3.3.3.2 Estimation strategy

The SFA requires specifying a function for the production frontier, $f(\boldsymbol{x}_i, \boldsymbol{w}_i; \boldsymbol{\beta}, \boldsymbol{\theta})$. Many production functions in the literature possess the quasi-concavity property suggested by producer theory and $f(\boldsymbol{x}_i, \boldsymbol{w}_i; \boldsymbol{\beta}, \boldsymbol{\theta})$ can take any of those forms. However, the Cobb-Douglas and translog production functions are two most commonly used specifications for the production frontier in SFA studies. The Cobb-Douglas functional form has the advantage that it is easier to estimate (because of fewer parameters) and the results are easier to interpret. However, it restricts output elasticities to be constant and elasticity of substitution between inputs to unity. The translog functional form is a generalization of the Cobb-Douglas functional form, and is more flexible because it does not impose restrictions on the elasticities; it allows elasticity of substitution to vary from point to point on the production function. However, the translog functional form requires estimation of many parameters (because of interaction terms) and the results are more difficult to interpret than in the Cob-Douglas. Because of its flexibility, we use the translog functional form, and test its suitability against the Cobb-Douglas specification. The model we estimate, i.e. equations (12a), (16) and (17), is specified below:

Frontier:

$$\ln y_i = \beta_0 + \sum_j \beta_j \, ln(x_{ij}) + \frac{1}{2}\sum_j \beta_{jj} \, ln(x_{ij})^2 + \sum_j \sum_{k \neq j} \beta_{jk} \, ln(x_{ij}) ln(x_{ik}) \tag{3.18}$$

$$+ \sum_s \theta_s \, w_{is} + \alpha T + \varepsilon_i$$

Inefficiency:

$$\mu_i = \delta_0 + \sum_l \delta_l \, z_{il}$$

$$\sigma_{ui}^2 = exp\ (\tau_0 + \sum_l \tau_l \, z_{il})$$

In the model (3.18), i, j, and s index maize farm, input, and environmental variable (soil fertility attribute or agronomic practice) respectively; y is output; $\boldsymbol{x}$ is the input vector; $\boldsymbol{w}$ is the vector of environmental variables (soil fertility conditions and agronomic practices); and $\boldsymbol{\beta}$ (also containing the intercept) and $\boldsymbol{\theta}$ are parameter vectors associated with inputs and their interaction terms and environmental variables, respectively. As explained earlier, Kenya's maize production is quite diverse with respect to environmental conditions. These conditions, such as soil fertility, are likely to affect maize yield while at the same time conditioning farmers' input use decisions. Failure to control for such conditions may result in biased production frontier and inefficiency estimates. We also include a dummy variable for year of the survey, T, to control for year differences and α is its associated parameter. $\varepsilon_i = v_i - u_i$ is the composed error term. The vector $\boldsymbol{z}$ in the inefficiency equations represents exogenous factors that influence technical efficiency and $(\boldsymbol{\delta}, \boldsymbol{\tau})$ are the corresponding parameter vectors to estimate.

Model (3.18) is estimated in a single step using maximum likelihood (ML) procedure as outlined earlier. The log-likelihood function is presented in Wang (2003) and technical efficiency estimates are obtained using Battese & Coelli (1988) method (equation (11)). The model is specified with and without including the environmental variables to understand the importance of controlling for environmental production conditions in agricultural productivity and efficiency estimation.

3.3.4 Data sources and variables

3.3.4.1 Data sources

The study uses household- and plot-level survey dataset on maize production in Kenya described in section 2.2.31. In addition, the study uses rainfall data obtained from the Climate Hazards group Infrared Precipitation with Stations (CHIRPS) dataset, developed in support for the United States Agency for International Development Famine Early Warning Systems Network (FEWS NET) (Funk et al., 2015). The dataset comprises high resolution (0.05°), daily, 5-day mean, and monthly precipitation that spans over 30 years (see Funk et al., 2015 for details). Household-level GPS coordinates were used to extract from the dataset daily precipitation data for growing periods of 2013/2014 and 2015/2016 main seasons of maize production. The daily precipitation data were summed over the growing period for each season to generate data on amount of rainfall for that season. All households within a sub-location had the same rainfall amount because they were within the same 0.05° x 0.05° grid.

3.3.4.2 Variables

Output variable

The frontier part of the stochastic frontier model specifies the relationship between output and inputs, soil fertility conditions and agronomic management practices. The output variable is maize yield per acre. Because intercropping maize with other crops, especially common beans, is the norm rather than an exception in Kenya, and because it is not possible to meaningfully apportion inputs applied on a plot to maize production, it is necessary to adjust for intercropping in calculating maize output on a plot. We create an output index for intercropped maize plots using the following method proposed by Liu & Myers (2009):

$$y_i = \frac{\sum_j y_{ij} P_j}{P_1} \quad , \tag{3.19}$$

where y_i is the output index on plot i, y_{ij} the yield of crop j on plot i and P_j the sub-county median price of crop j, computed from producer prices reported by the sample households that sold crop j. Crop 1 is maize, so that P_1 is the price of maize.

Input variables

The inputs are inorganic fertilizers, maize seed, labor, and mechanization[3]. Because maize farmers often use a wide range of fertilizers and in varied amounts and because different fertilizers have different amounts of macronutrients, it is reasonable to use amounts of macronutrients rather than the quantity of fertilizer applied in the production frontier equation. Compound inorganic

[3] While farmers also use organic materials such as manure and compost, we consider their application among agronomic practices because farmers apply them with a long-term view of improving the fertility condition of the soil rather than for production in a growing season.

fertilizers often have three macronutrients - nitrogen (N), phosphorus (P) and potassium (K). The nutrient composition of the compound fertilizers is identified by percentages of nitrogen (N), phosphate (P_2O_5) and potash (K_2O) (Maguire, Mark, & Flowers, 2009), which, as a standard practice internationally, are normally labeled on fertilizer packaging (bags). Such labeling helps in determining the amount of the macronutrients in each quantity of fertilizer and is, therefore, useful to decisions about appropriate fertilizers and quantities to apply. For example, the most commonly used planting fertilizer for maize in Kenya is diammonium phosphate (DAP), which has the composition (formula) (18:46:0). This means that any amount of DAP has 18% N, 46% P_2O_5 and 0% K_2O. To get the amount of each macronutrient applied on a plot, first we need to know the percentage of phosphorus (P) in the available phosphate (P_2O_5) and potassium (K) in the water-soluble potash (K_2O), then multiply these with the respective composition percentages and eventually by the amount of fertilizer applied. P constitutes approximately 43.6% of P_2O_5 while K constitutes 83% of K_2O. Continuing with the DAP example, it means that any amount of DAP contains 18% N, about 20% (0.436*46) P and 0% (0.83*0) K. The remaining 62% is filler material. Another fertilizer among many that farmers use is NPK (23:23:23). Any amount of this fertilizer contains 23% N, 10% (0.436*23) P and 19% (0.83*23) K, and the remaining material is filler. We compute the amount of the macronutrients from all the fertilizers that farmers reported for their largest maize plots. Because of the fertilizer formulations, farmers apply the macronutrients in fixed proportions when they use inorganic fertilizers. There is thus likely to be high collinearity in the amount of N, P and K applied on a plot. Indeed, data shows a high correlation coefficient of 0.74 between N and P applied to maize plots. Therefore, we use the amount of N as a measure of the amount of fertilizer applied to a plot. The use of N to represent fertilizer amount is reasonable since nitrogen is the most important and often limiting nutrient in the soil.

Maize seed is measured as quantity (in kg) of seed planted per acre. Because not all farmers planted improved seed varieties (i.e. hybrid or open pollinated variety (OPV)), we control for the effects of improved seed variety using a dummy variable.

We measure labor as person-days per acre. It constitutes family labor by adult household members, hired labor and any other labor that is neither family nor hired. Data on children's labor is available for only 2014 and is excluded from the analysis for that reason and because of lack of a standard factor for converting it into adult-equivalent labor. As in Liu & Myers (2009), only labor on pre-harvest activities is included in the analysis because harvest and post-harvest activities are not expected to have significant effects on maize yield.

The effect of mechanization is captured through a dummy variable for tractor and/or oxen use in land preparation and/or planting, the two pre-harvest activities for which mechanized operation is most common and data were collected.

Although we measure output and inputs in terms of amount per unit area, we also include plot size explicitly in the frontier function so as not to impose constant returns to scale of the production technology in land.

Soil fertility conditions

As explained earlier, a major contribution of this study is in controlling for the effects of soil fertility conditions and agronomic practices in evaluating technical efficiency. This overcomes potential omitted variable bias to which many studies that do not do so may be prone. Soil fertility conditions affect the responsiveness of crops to inputs, especially fertilizers, and so it is important to control for variability in these conditions in a production function estimation. Because of data limitations, some studies have used geographical location dummies and average soil fertility conditions at larger geographical scales (e.g. Liu & Myers, 2009; Sheahan, 2011). However, soil fertility conditions can dramatically vary even within a very small geographical area, so using average soil conditions for large geographical areas to represent soil fertility on cultivated plots certainly does not accurately capture the variability in soil fertility across the plots. This study overcomes that limitation by using plot-level measures of soil physical and chemical properties to control for the effects of soil fertility status on maize yield.

Doran & Parkin (1994) propose a selection of physical, chemical and biological properties of soil that can be useful as indicators for assessment of soil fertility. Based on the soil data available for this study, we control for soil fertility using measured soil physical and chemical properties. Soil texture, a physical property, is a measure of relative proportion (by weight) of clay, silt and sand particles in a soil. Texture influences many soil properties, including drainage, water holding capacity, erodibility, organic matter content and aeration. We measure soil texture in terms of sand content (i.e. percentage sand). Soils higher in sand content have lower water holding capacity and organic matter content, which is important for nutrient storage and availability to plants. We expect maize yield to vary inversely with sand content in the soil.

Among the soil chemical properties proposed by Doran & Parkin (1994) as indicators that can be used to assess soil fertility include total organic carbon, total nitrogen, soil pH, electrical conductivity and extractable (or plant-available) macronutrients (P, K). Total carbon and total nitrogen are indicators of soil fertility and potential crop productivity of the soil and are highly correlated. We therefore use amount of total carbon in the soil, measured in percentage. Soil pH is a measure of soil acidity, which affects biological and chemical activity in the soil. Lower pH levels indicate greater acidity while higher values indicate alkalinity. Most crops require near neutral levels of pH to grow best. Electrical conductivity of soil is a measure of the amount of soluble (salt) ions in the soil. It varies depending on a host of soil properties, including texture, water holding capacity, organic matter content and cation exchange capacity. Electrical conductivity is thus strongly correlated with many soil properties, particularly texture (Grisso et al, 2009), and, for this reason, we do not include it in the model. We include plant-available phosphorus among the macronutrients because it is particularly sensitive to soil acidity.

Agronomic practices

Agronomic practices refer to farm management actions that farmers take to improve soil fertility, enhance utilization of agricultural production resources and improve the environment. While such practices are many and cut across different aspects of the farm, this study concentrates on the group of practices often targeted at improving soil fertility to increase crop production. A dummy variable for use of manure (animal or green) or compost on the plot is included to capture the effects on maize yield of management practices that improve soil fertility. Addition of manure or compost increases soil organic matter, which acts as a reservoir of nutrients in forms that are available to plants and improves soil structure, maintains soil tilth and minimizes erosion (Bot & Benites, 2005). Also included are dummy variables for use of crop residues and intercropping maize with legumes (e.g. common beans, cowpeas, pigeon peas). Crop residues, upon decomposition, add organic matter into the soil while legumes fix nitrogen in the soil hence improves soil fertility.

To control for the effects of rainfall on maize yield, we include two variables; rainfall amount and rainfall stress during the growing season. We measure stress as the fraction of 20-day periods that received less than 40 mm of rain during the growing season. Rainfall stress is important because it captures rainfall distribution over the growing period. Maize crop demand for water varies with growth stages and it is important that at every stage there is sufficient supply of water for better yield potential. We control for other differences in environmental, policy and market conditions through county dummy variables. This is very important because since 2012 Kenya's department of agriculture, livestock and fisheries has been under the management and policy directives of respective county governments, which operate independently in terms of policies and programmes they prioritize and implement.

Exogenous factors influencing efficiency

Kagin, Taylor, & Yúnez-Naude (2015) note that theory offers little guidance on what explains technical efficiency and suggest that its analysis is primarily an empirical endeavor. However, literature has established existence of statistical relationship between technical efficiency and a range of contextual characteristics under which farms operate. We thus follow the existing literature (e.g. Yang et al., 2016; Kumbhakar et al., 2014; Liu & Myers, 2009; Wang, 2002; Battese & Coelli, 1995; Kumbhakar et al., 1991) in selecting the variables that go into the inefficiency part of the stochastic frontier model. Factors that affect efficiency are primarily those that often influence a farmer's management capacity of the production process. It is possible that different groups of farmers can use the same type and amounts of inputs and operate under the same environmental conditions but produce different output levels. Such differences can be explained by factors that affect how they manage their production process.

We use education to control for human capital and measure it as the education level of the plot manager. Kumbhakar et al. (1991), Battese & Coelli (1995), Liu & Myers (2009) and Yang et al., (2016) found a positive relationship between education and technical efficiency. Also related to human capital are skills acquired through experience in farming and so we include as a measure of experience the number of years the plot manager's household has been in farming. To control for the effects of gender on technical efficiency, we include a dummy variable for female plot manager. It is widely acknowledged that women relative to men are generally disadvantaged in terms of agricultural productive resources, including human capital.

A farmer's incentives to invest in long run productivity improvement can be affected by the attributes of their land, including tenure security (Besley, 1995). Location of the cultivated plot can also affect its management, with plots near homesteads receiving greater attention because they are easily reachable and monitoring them should be more convenient. Research has also found inverse farm size-productivity relationship and farm size-technical efficiency relationship (e.g. Benjamin, 1995; Kagin, Taylor, & Yúnez-Naude, 2016). These show that it is important to control for the effects of farmers' land attributes in inefficiency estimation. We use a dummy variable for plot ownership to control for effects of tenure security. Two measures are used to control for the effects of land size – household total landholding and the size of the maize plot. We also include distance from the household's dwelling to the maize plot.

Farmers' perceptions and/or knowledge about the fertility condition of soil on their farms can partly influence how they manage their agricultural production on the farm. However, such perceptions may not necessarily reflect the actual soil fertility condition and thus may result in management practices that mismatch the needs of the soil. If this happens, it may have undesirable influence on technical efficiency. We control for the effects of deviation of farmers' perceptions about the fertility conditions of their soils from measured fertility as determined from soil test results. Farmers were asked to rate the fertility of soil on their maize plots on a Likert scale of 1 (very infertile) to 4 (very fertile) based on their own perceptions. We reduce the scale to two categories - infertile (combining 1 and 2) and fertile (combining 3 and 4) - because of very low responses on the two extreme categories (1 and 4). From the soil test data, we use three chemical properties of soil – total carbon (C), total nitrogen (N) and pH - and their threshold values as obtained from recommendations by the Kenya government on critical levels of various soil nutrients and pH for maize growth, to determine whether a soil is infertile or fertile. A plot is fertile if the soil has total C$\geq$ 2.7%, N$\geq$ 0.2% and $5.5 \leq \text{pH} \leq 7.0$. We compare a farmer's rating (fertile or infertile) against the rating using these threshold values, and construct a binary variable for farmer perception deviation that takes the value of 1 if the farmer's rating differs from the rating using threshold values and 0 otherwise.

We control for the effects of extension services using distance to extension service provider. While suitable variables to capture effects of extension services include number of visits to extension service provider and whether a farmer received extension advice, such information is, unfortunately, lacking in the data.

We exclude from the analysis maize plots with any of the following characteristics: 1) intercropped plots for which the share of maize in total value of crops output is less than 20% because the focus of the study is on maize; 2) less than 0.2 acres in size because of potential for measuring error in computing rates of input use; 3) labor use rate of less than 3 or more than 3000 person-hours per acre; and 4) maize yield less than 80kg/acre. Ultimately, the analysis uses 1102 pooled plot observations in 621 households. Summary statistics of the variables is presented in Table 3.1.

Table 3.1: Variable description and summary statistics (N=1102)

Variable	Unit of measure	Mean	SD	Min	Median	Max
Output						
Maize yield index	Kg/acre	1256.15	842.86	81.14	1089.32	7800.00
Inputs						
Nitrogen amount in fertilizer applied	Kg/acre	19.57	16.48	0.00	16.26	109.00
Maize seed	Kg/acre	9.50	2.81	2.00	10.00	20.00
Pre-harvest labor	Person-hours/acre	229.05	217.90	3.00	172.00	2484.00
Improved maize seed	1=yes, 0=no	0.94	0.24	0.00	1.00	1.00
Mechanized land preparation	1=yes, 0=no	0.54	0.50	0.00	1.00	1.00
Agronomic practices						
Manure/composed use	1=yes, 0=no	0.44	0.50	0.00	0.00	1.00
No. of crops	Number	2.18	1.10	1.00	2.00	9.00
Maize intercrop with legume	1=yes, 0=no	0.72	0.45	0.00	1.00	1.00
Use of crop residue	1=yes, 0=no	0.42	0.49	0.00	0.00	1.00
Soil fertility conditions						
Total organic carbon	% (value)	2.12	0.86	0.22	2.00	4.93
Plant available phosphorus	Parts per million (ppm)	17.77	18.90	0.20	12.40	168.00
Soil pH	Value	5.67	0.51	4.40	5.62	8.03
Soil Texture	% sand	48.13	18.83	7.55	50.00	88.10
Main season rainfall	mm	1015.36	692.67	263.08	843.72	3294.94
Moisture stress (Fraction of 20-day periods with <40 mm of rainfall)	(0-1)	0.21	0.16	0.00	0.13	0.60
Inefficiency predictors						
Education of plot manager	years	7.97	3.88	0.00	8.00	17.00
Gender of plot manager	1=female, 0=male	0.35	0.48	0.00	0.00	1.00
Experience of household in farming	years	27.51	16.07	1.00	25.00	87.00
Farmer perception of plot fertility	1=fertile, 0=infertile	0.59	0.49	0.00	1.00	1.00
Consistency of farmer perception and measured plot fertility	1=consistent, 0=inconsistent	0.47	0.50	0.00	0.00	1.00
Plot size	acres	1.15	2.08	0.20	0.50	30.00
Household land holding	acres	3.20	5.40	0.00	1.50	49.00
Distance to plot	Walking time (minutes)	4.75	10.65	0.00	2.00	180.00
Household owns plot	1=yes, 0=no	0.89	0.31	0.00	1.00	1.00
Distance to extension service	km	5.75	4.98	0.10	5.00	40.00

3.4 Results

Model diagnostic test results are discussed first, followed by maximum likelihood estimation results of the stochastic frontier production function, and, finally, technical efficiency estimates and implications.

3.4.1 Model diagnostic test results

3.4.1.1 Likelihood ratio test of presence of technical inefficiency

We test for the presence of technical inefficiency using the likelihood ratio test procedure as outlined in Kumbhakar, Wang, & Horncastle (2015). We compare the log likelihood ratio of OLS regression against that of maximum likelihood estimation of the Wang (2002) stochastic frontier model, in which both the pre-truncated mean and variance of technical inefficiency are functions of exogenous variables. We conduct the test with and without environmental variables (soil fertility conditions and agronomic practices) in the estimation equations, and use the flexible form of the translog specification for the production frontier. The likelihood ratio (LR) statistic is computed as -2[L(H0)-L(H1)]. L(H0) is the log-likelihood value of the OLS regression (restricted model) while L(H1) is that of the stochastic frontier (unrestricted model). The degree of freedom is the number of restrictions, which in our case is 22; i.e. 11 parameters in each of the pre-truncated mean and variance of the inefficiency term. The LR statistic is compared to the critical values of the mixed chi-square distribution. Results show that the LR statistic is 112.509 when environmental variables are included and 116.167 when environmental variables are not included. Both are larger than the critical value of 39.664 at 1% significance level. We thus reject the null hypothesis of no technical inefficiency irrespective of whether environmental variables are included in the estimation. The OLS regression results are in Table A3.1 in the Appendix.

3.4.1.2 Production frontier functional form test: translog vs Cobb-Douglas

We conduct a likelihood ratio test comparing the suitability of the translog functional form (unrestricted model) against the Cobb-Douglas functional form (restricted model) for the production frontier. We conduct the test separately for the model specification with and without environmental variables. The number of restrictions is equal to the quadratic and interaction terms in the translog functional form, which in this case is six. In the model specification with environmental variables, the LR statistic of the test is 41.70 with $p<0.001$, while in the specification without environmental variables the LR test statistic is 33.48 with $p<0.001$. Therefore, the translog functional form fits the data better than the Cobb-Douglas irrespective of whether environmental variables are included in the production frontier equation. Maximum likelihood estimation results of the Cobb-Douglas production frontier (with and without environmental variables) are in Table A3.2 in the Appendix. Subsequent tests and analyses proceed with the flexible translog functional form specification for the production frontier.

3.4.1.3 Stochastic frontier model test

This study uses the Wang (2002) model in which both the pre-truncated mean and variance of technical inefficiency are modelled as functions of exogenous variables, thus combining features of the model of Kumbhakar, Ghosh, & McGuckin (1991), Huang & Liu (1994) and Battese & Coelli (1995) (KGMHLBC) and that of Caudill & Ford (1993), Caudill, Ford, & Gropper (1995) and Hadri (1999) (CFCFGH) models. For completeness, we test the suitability of this model against the KGMHLBC and CFCFGH models. We also test its suitability against the truncated-normal model of Stevenson (1980) in which exogenous variables do not influence the distribution of inefficiency. We use LR test where the Wang (2002) model is the unrestricted model and the others the restricted models. We conduct the tests for the specifications with and without environmental variables. Test results are shown in Table 3.2 (Maximum likelihood estimates of the restricted models are presented in Table A3.3 in the Appendix). Clearly, the Wang (2002) model best fits the data. Further, rejection of the Stevenson model in favor of the Wang (2002) model affirms that inefficiency is affected by the exogenous factors as specified in the Wang (2002) model. Subsequent analyses and results are based on the Wang (2002) model.

Table 3.2: Likelihood ratio test results of the Wang (2002) model against KGMHLBC, CFCFGH and Stevenson models

Models	No. of restrictions	With environmental variables		With environmental variables		5% critical value	Decision
		LR statistic	p-value	LR statistic	p-value		
Wang vs KGMHLBC	10	33.73	0.0002	25.58	0.0044	18.307	Reject KGMHLBC
Wang vs CFCFGH	10	31.04	0.0006	28.68	0.0014	18.307	Reject CFCFGH
Wang vs Stevenson	20	97.89	0.0000	97.51	0.0000	31.410	Reject Stevenson

3.4.2 Stochastic production frontier estimation results

The joint maximum likelihood estimates of the frontier part of model (3.18), specified in flexible translog functional form, are presented in Table 3.3. The estimation was conducted with and without environmental variables and both sets of results are presented. Likelihood ratio test shows that the model specification with environmental variables fits the data better; the LR statistic (with 13 degrees of freedom) is 139.95 with $p<0.001$. Under the specification with environmental variables, maize yield is positively and significantly correlated with fertilizer (N) and seed. The positive and significant coefficients of the first and second order terms of the fertilizer variable indicates that for the entire range of the data, increased use of nitrogen fertilizer would, on average, have increasing effects on maize yield. Although weakly significant, the negative sign on the second order term of fertilizer and labor indicate that the two inputs are substitutes. As expected, maize yield is positively correlated with use of improved seed varieties. In addition, yield has a positive association with mechanized land preparation. There is an inverse relationship between maize yield and plot size, which is consistent with findings in majority of empirical studies on farm size-productivity relationship (see Barrett, 1996; Barrett, Bellemare, & Hou, 2010; Carletto, Savastano, & Zezza, 2013). The positive association of maize yield with the number of crops on a plot indicates that on average, intercropped maize plots would have higher yield than monocropped plots.

The coefficient estimates of the production frontier without environmental variables included are qualitatively similar to those in the specification with environmental variables. The only major differences are that in the specification without environmental variables, the coefficient on the first order term for fertilizer is not statistically significant at any reasonable significance level, while the coefficient on improved seed is strongly significant.

Regarding environmental variables, manure/compost use, higher total organic carbon in the soil and soil pH are each associated with higher maize yield. Holding other things constant, manure/compost use is associated with approximately 15% higher maize yield, on average, while the coefficient on soil pH indicates that increasing the pH by 10% is associated with 1.7% increase in maize yield, suggesting that the sample maize plots are acidic, on average. Indeed 39% of the plots in the sample had soil pH values below the recommended minimum of 5.5, and 52% had values below the sample average of 5.7. Therefore, high soil acidity on maize farms in Kenya is an issue that needs to be addressed.

Table 3.3: Stochastic production frontier estimates

Response variable: Log of maize yield	With environmental variables		Without environmental variables	
	Coefficient	Std. error	Coefficient	Std. error
Explanatory Variables				
Log of fertilizer (lnN)	0.249*	0.149	0.178	0.162
Log of seed (lnseed)	0.814**	0.384	0.840**	0.415
Log of labor (lnlabour)	-0.0549	0.173	0.0271	0.184
lnN x lnN	0.128***	0.0250	0.145***	0.0261
lnseed x lnseed	-0.221	0.186	-0.104	0.199
lnlabour x lnlabour	0.0169	0.0229	0.00820	0.0234
lnN x lnseed	-0.145	0.0991	-0.126	0.105
lnN x lnlabour	-0.0636*	0.0366	-0.0760*	0.0403
lnseed x lnlabour	0.0944	0.118	0.0409	0.126
Improved seed (1=yes)	0.127*	0.0762	0.292***	0.0793
Mechanized land preparation (1=yes)	0.147***	0.0511	0.252***	0.0478
Log of plot size	-0.149***	0.0285	-0.120***	0.0335
No. of crops on plot	0.117***	0.0203	0.0825***	0.0174
Year dummy (1=2016)	0.177***	0.0555	0.282***	0.0369
Manure/compost use (1=yes)	0.142***	0.0363		
Maize-legume intercrop (1=yes)	-0.0711	0.0497		
Crops residue use (1=yes)	0.0367	0.0351		
Total organic carbon	0.0460*	0.0276		
Phosphorus	-0.00162	0.00101		
Soil pH	0.171***	0.0357		
Sand content	-0.00181	0.00159		
Rainfall	0.0000409	0.0000369		
Moisture stress	-0.217	0.136		
Transnzoia county dummy[a]	-0.137**	0.0585		
Kakamega county dummy[a]	-0.341***	0.0717		
Kisii county dummy[a]	-0.697***	0.0868		
Machakos county dummy[a]	-0.377***	0.0833		
Constant	3.933***	0.732	4.310***	0.771
σ_v^2	-1.623***	0.101	-1.563***	0.132
Log-likelihood	-863.1077		-933.0820	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: [a]Uasin Gishu is the comparison county

Coefficients on the other environmental variables – plant available phosphorus, soil texture (sand content), rainfall and moisture stress – are not statistically significant although all but plant available phosphorus have expected signs. The unexpected sign of the coefficient on plant available phosphorus is similar to results found by Ekbom et al (2013) in their study of effects of soil capital on maize productivity in central Kenya. They attribute their result in part to high soil acidity, which, as explained earlier, has the effect of causing phosphorus to form insoluble compounds, which makes it unavailable to plants. In highly acidic soils, tests may indicate high levels of phosphorus but phosphorus deficiency in plants may occur and reduce output.

The input coefficients in Table 3.3 are not quite informative by themselves about the relationship between maize yield and the inputs because the interaction terms confound such relationships. We, therefore, computed output elasticities of the inputs, which have straightforward and meaningful interpretation. The elasticity estimates were computed for each plot and the means are presented in Table 3.4. Output elasticity of N is particularly of interest, since fertilizer often seems to be the most limiting input in maize production and which has drawn much interest among policy makers and researchers in Kenya and the region. The average elasticity of N is 0.262 when environmental variables are controlled for. This estimate is statistically larger than the 0.226 average estimate without controlling for environmental variables (a paired sample t-test of difference in mean has t= 59.815). In both model specifications, the elasticities vary considerably across the counties, with Uasin Gishu having the highest and Machakos the lowest estimates. The elasticity estimates for N compare well with those obtained by other researchers using data from Kenya. For example, Mghenyi (2015) found output elasticity of N of 0.18 while Liu & Myers (2009) found it to be 0.224 among hybrid maize seed users and 0.209 among local seed users. These studies, however, did not control for environmental variables.

Table 3.4: Maize output elasticity with respect to inputs (mean)

County	With environmental variables			Without environmental variables		
	Fertilizer (N)	Seed	Labor	Fertilizer (N)	Seed	Labor
Uasin Gishu	0.317	0.294	0.0336	0.293	0.504	-0.00154
	(0.00887)	(0.00706)	(0.00262)	(0.0101)	(0.00484)	(0.00260)
Trans Nzoia	0.290	0.330	0.0463	0.260	0.521	0.00594
	(0.00891)	(0.00661)	(0.00243)	(0.0101)	(0.00482)	(0.00265)
Kakamega	0.295	0.362	0.0431	0.263	0.536	0.00425
	(0.00765)	(0.00688)	(0.00210)	(0.00868)	(0.00473)	(0.00228)
Kisii	0.241	0.421	0.0593	0.200	0.572	0.0198
	(0.00678)	(0.00684)	(0.00197)	(0.00767)	(0.00451)	(0.00202)
Machakos	0.174	0.470	0.0745	0.123	0.611	0.0399
	(0.00839)	(0.00781)	(0.00223)	(0.00952)	(0.00529)	(0.00251)
Overall	0.262	0.376	0.0517	0.226	0.550	0.0141
	(0.00401)	(0.00372)	(0.00112)	(0.00458)	(0.00248)	(0.00119)

Standard errors in parentheses
Note: Elasticities are computed at plot level

The average elasticities of seed are 0.376 and 0.550, respectively, with and without controlling for environmental variables, and vary across the counties, with the estimates highest in Machakos and lowest in Uasin Gishu, an exactly opposite pattern to the one for the elasticity of N. Liu & Myers (2009) estimated elasticities of seed to be 0.336 and 0.293 among hybrid and local seed users, respectively, while, surprisingly, Mghenyi (2015) found a negative elasticity of -0.3 for seed.

Output elasticity of labor averages 0.0517 and 0.0141 for the specification with and without environmental variables, respectively, and differs across counties. These estimates are quite different from those obtained by Liu & Myers (2009) (0.177 and 0.300 among hybrid and local seed users, respectively) but are close to Mghenyi's (2015) estimate (0.04).

The elasticity estimates point to the importance of controlling for environmental variables in agricultural production function estimation. Controlling for soil fertility conditions is especially important for accurate estimation of the effects of fertilizer on yield. This point will be clearer below when we discuss technical efficiency estimates and, subsequently, marginal physical product of the inputs.

3.4.3 Technical efficiency estimation results

The main aim of this study is to estimate technical efficiency of maize farmers and identify factors responsible for variation of technical efficiency across farms, and to understand how important it is to account for environmental conditions in the analysis. We first present results on the marginal effects of exogenous factors on technical inefficiency from the joint maximum likelihood estimation of model (3.18). Next, we discuss technical efficiency estimates and how these vary across regions (counties), and the difference in the estimates when we do and not account for environmental conditions. Marginal products of the variable inputs are also computed and compared across regions to understand regional variations in potential output response to the variable input use.

3.4.3.1 Determinants of technical (in)efficiency

Marginal effects of exogenous variables on the unconditional mean and variance of inefficiency are presented in Table 3.5. The marginal effect on unconditional mean measures how expected inefficiency changes when the exogenous variable in question increases, while marginal effect on unconditional variance measures the effect of an increase in the exogenous variable on uncertainty of technical inefficiency (Wang, 2002; Kumbhakar et al., 2015). It is important to note that for continuous variables, a negative sign of the marginal effect on expected inefficiency implies an increasing effect on the level of technical efficiency for a positive change in the variable, and vice versa. For binary variables, a negative sign of the marginal effect on the expected inefficiency implies that technical efficiency is, on average, higher compared to the base (omitted) category.

Table 3.5: Marginal effects of exogenous determinants of technical inefficiency (sample means)

Variables	With environmental variables		Without environmental variables	
	Marginal effect	Bootstrap std. error	Marginal effect	Bootstrap std. error
Mean function (μ)				
Plot manager education (yrs)	-0.0213***	0.000340	-0.0144***	0.000170
Plot manager is female (1=yes)	-0.0546***	0.00592	-0.0221***	0.00698
Soil fertility perception (1=fertile)	-0.191***	0.00716	-0.236***	0.00295
Farmer perception consistent with soil test (1=yes)	-0.0286***	0.00643	-0.0502***	0.00322
Plot size (acres)	-0.0478***	0.000953	-0.0812***	0.00193
HH landholding (acres)	0.000746***	0.000190	-0.00659***	0.000266
Walking time to plot (mins)	-0.00240***	0.000151	-0.00220***	0.0000797
Plot owned (1=yes)	0.328***	0.0171	0.0651***	0.00253
Distance to extension (km)	0.000207*	0.000120	0.00154***	0.000130
Farming experience (yrs)	-0.00193***	0.000188	-0.00137***	0.000138
Variance function (σ_u^2)				
Plot manager education (yrs)	-0.0103***	0.000223	-0.00630***	0.0000938
Plot manager is female (1=yes)	-0.0636***	0.00283	-0.0754***	0.00242
Soil fertility perception (1=fertile)	-0.144***	0.00490	-0.144***	0.00275
Farmer perception consistent with soil test (1=yes)	-0.0523***	0.00272	-0.0554***	0.00143
Plot size (acres)	-0.0218***	0.000444	-0.0256***	0.000362
HH landholding (acres)	-0.000672***	0.0000624	-0.00581***	0.000137
Walking time to plot (mins)	-0.00216***	0.0000841	-0.00186***	0.0000429
Plot owned (1=yes)	0.0849***	0.00360	0.0112***	0.000416
Distance to extension (km)	-0.000572***	0.0000424	-0.000379***	0.0000305
Farming experience (yrs)	0.00000600	0.0000500	0.000549***	0.0000344
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: A negative sign of the marginal effect on the mean function implies a contribution to efficiency while a positive sign is a contribution to inefficiency for a positive change in the variable.

We observe qualitatively similar (in terms of signs) marginal effects of the exogenous variables between the model specification with and without environmental variables. The only exception is in the landholding variable in the mean function for which the effects are positive in the specification with but negative in that without environmental variables. Quantitatively, however, the marginal effects are quite different between the two model specifications but with no distinct pattern, although the magnitudes of the effects are generally small in both specifications. With respect to the unconditional mean and variance in each of the two model specifications, the signs of the marginal effects of nearly all the exogenous variables are the same. We focus the discussion on the marginal effects in the model specification with environmental variables.

On average, both the level and uncertainty of technical inefficiency tend to reduce with increased education level of plot manager, plot size and distance to plot, and when plot manager is female, while they tend to be higher when the plot is owned (rather than rented). Increased farming experience has a decreasing effect on the level of technical inefficiency. Increased landholding and distance to extension service provider have increasing effects on the level but decreasing effects on uncertainty of technical inefficiency. These effects are largely as expected, except for the effects of gender of plot manager, distance to plot, plot ownership and landholding. The unexpected effect of distance to plot and plot ownership may reflect management efforts farmers put on rented plots, which the data show are located on average three times farther from the homestead than are owned plots. Paying rent on a plot may incentivize a farmer to apply management practices for better crop yield to increase returns to investment. The effects of gender of plot manager contradicts findings by other studies that have shown that female-headed households are more technically inefficient than their male-headed counterparts (e.g. Liu & Myers, 2009).

Concerning farmer perception about soil fertility status, the first notable result is that farmers that view their plots as fertile have, on average, 19% lower technical inefficiency than do their counterparts that consider their plots infertile. They also have less uncertainty of inefficiency. The second is that those whose perceptions of the fertility status of their plots are consistent with that determined from soil test have 2.9% lower technical inefficiency than do those with inconsistent view and they also have less uncertainty of inefficiency, although these are only significant at 10% level. These effects indicate the importance of farmer perceptions about the fertility conditions of their soils and underscores the need for farmer access to information that can enhance their knowledge about correct soil fertility conditions on their farms.

3.4.3.2 Technical efficiency estimates

Technical efficiency estimates were computed using Battese & Coelli (1988) method: $E(exp(-u_i)|\varepsilon_i)$. Figure 3.4 shows kernel densities of technical efficiency for model specification with and without environmental variables. Clearly, technical efficiency estimates based on the model specification with environmental variables are higher compared to those based on the specification without environmental variables. The mode of technical efficiency for the specification with environmental variables is around 0.83 while it is around 0.79 for the specification without environmental variables. The distribution of efficiency estimates is left-skewed in both specifications. Descriptive statistics in Table 3.6 show that technical efficiency ranges from 0.12 to 0.99, with the mean at 0.76, for the specification with environmental variables, while without environmental variables the estimates range from 0.19 to 0.98, with mean at 0.70. The difference in the two means is statistically significant ($t=9.021$, $p<0.001$). This is evidence that failure to account for the effects of environmental variables in SFA underestimates technical efficiency, a finding that is consistent with Sherlund et al (2002) in a study of technical efficiency in rice production in Ivory Coast.

Figure 3.4: Kernel density of technical efficiency estimates

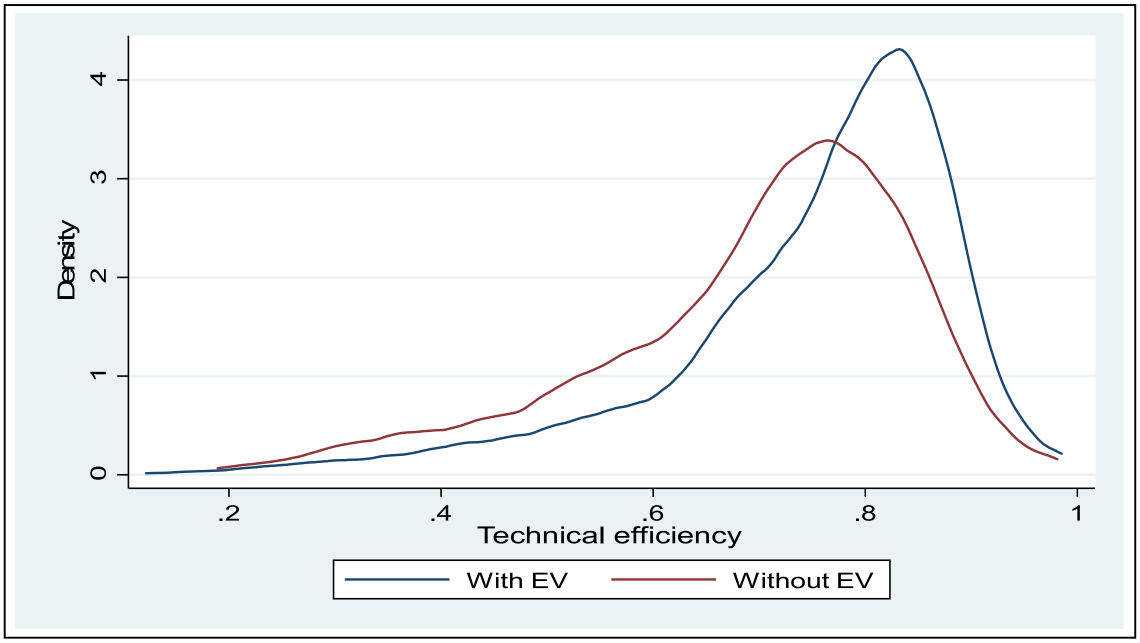


Table 3.6: Technical efficiency estimates (mean)

County	With environmental variables				Without environmental variables			
	Mean	Median	Minimum	Maximum	Mean	Median	Minimum	Maximum
Uasin Gishu	0.751	0.798	0.123	0.985	0.755	0.784	0.217	0.978
Trans Nzoia	0.766	0.783	0.305	0.983	0.738	0.754	0.314	0.981
Kakamega	0.750	0.785	0.248	0.984	0.669	0.707	0.218	0.893
Kisii	0.740	0.769	0.303	0.952	0.629	0.659	0.191	0.895
Machakos	0.747	0.797	0.169	0.917	0.683	0.727	0.205	0.892
Overall	0.751	0.789	0.123	0.985	0.696	0.730	0.191	0.981

Note: Technical efficiency estimates are computed using Battese & Coelli (1988) method: $E(exp(-u_i)|\varepsilon_i)$

While the pattern of the difference in technical efficiency estimates between the two model specifications remains qualitatively the same within each county, except for Uasin Gishu where there appears to be hardly any difference, it is notable that controlling for environmental variables makes significant difference in the average technical efficiency estimates in Kisii, Kakamega and Machakos. Wald test shows that differences in average technical efficiency across the five counties is statistically significant only in the specification without environmental variables, suggesting that after controlling for environmental variables, technical efficiency in maize production is pretty much the same across different regions of Kenya. These results affirm that failure to account for differences in environmental conditions can give misleading results on the level of technical efficiency and its variation across regions.

Using equation (4) and estimated technical efficiency levels, we computed frontier maize yield and generated inefficiency-induced foregone output. Results show that inefficiency-induced maize yield loss in model specifications with and without environmental variables is 0.34 and 0.45 tonnes per acre, respectively (Table 3.7). Because technical efficiency is underestimated in the specification without environmental variables, the computed foregone yield appears to be higher. Nevertheless, the yield loss of one third of a tonne/acre is substantial, given that the average reported (actual) yield is just 1.3 tonnes/acre. Across the counties, yield losses due to inefficiency are highest in Uasin Gishu and Trans Nzoia and lowest in Kisii (in the model specification with environmental variables).

Table 3.7: Actual, frontier and foregone output (mean)

County	Actual output (kg/acre)	With environmental variables		Without environmental variables	
		Frontier output (kg/acre)	Foregone output (kg/acre)	Frontier output (kg/acre)	Foregone output (kg/acre)
Uasin Gishu	1497	1921	424	1912	416
Trans Nzoia	1563	1985	423	2042	480
Kakamega	1213	1548	336	1728	515
Kisii	846	1087	241	1272	425
Machakos	1128	1415	287	1525	398
Overall	1256	1599	343	1702	446

Note: Foregone output computed at the mean of actual and frontier output

Results on average marginal product (MP) of fertilizer (N), seed and labor are presented in Table 3.8. The marginal products are computed for each plot from the elasticity estimates. Overall, the MP of N, seed and labor are all positive, on average, suggesting that at the current average application rates, farmers would realize positive maize yield returns if they increased their rate of use of each of the inputs. There is significant variation in the average MP of inputs between the two model specifications. Focusing on the specification with environmental variables, the MP of fertilizer is 18.63 kg/acre, on average, and varies considerably across the counties. The MP of N is particularly high in Trans Nzoia and Uasin Gishu but below the total average in Kisii and Kakamega. The result on the MP of N is pretty close to those found by Mghenyi (2015) (14.81 kg/kg) and Marenya & Barrett (2009) (17.64 kg/kg). It is worth noting that Marenya & Barrett (2009) controlled for some soil properties and other environmental production conditions while Mghenyi (2015) did not. It is interesting to note that Mghenyi's (2015) estimate is only slightly larger than our estimate in the specification without environmental variables (13.01 kg/kg). Therefore, the fact that the estimate by Mghenyi (2015) is much lower than ours, which is larger than Marenya & Barrett (2009)'s estimate, affirms that omission of environmental variables from agricultural productivity analysis can lead to biased results.

The MP of seed averages 44.9 kg/kg while that of labor is quite low (0.15 kg/person-hour). These also vary across the counties, with Machakos having the highest MP of seed and Kisii the lowest. Trans Nzoia has the highest MP of labor while Uasin Gishu has the lowest and is negative. Our estimate of the MP of seed is incomparable to Mghenyi (2015) who found a large negative estimate (-49.70 kg/person-hour) (Marenya & Barrett (2009) did not include seed among the inputs in their estimation). However, our estimate of MP of labor, although above that of Marenya & Barrett 's (2009) estimate of 0.08 kg/man-day, is equally quite low.

Table 3.8: Marginal product of inputs

County	With environmental variables			Without environmental variables		
	Fertilizer (N) (kg/kg)	Seed (kg/kg)	Labor (kg/person-hour)	Fertilizer (N) (kg/kg)	Seed (kg/kg)	Labor (kg/person-hour)
Uasin Gishu	20.11	39.14	-0.356	12.97	62.36	-1.478
	(2.237)	(1.417)	(0.212)	(4.994)	(1.363)	(0.319)
Trans Nzoia	21.65	46.36	0.365	16.92	68.87	-0.134
	(0.569)	(1.480)	(0.0207)	(0.596)	(1.507)	(0.0331)
Kakamega	17.02	44.11	0.242	15.91	70.26	-0.0501
	(0.376)	(1.443)	(0.0150)	(0.336)	(1.790)	(0.0180)
Kisii	14.55	37.51	0.177	11.90	57.97	0.0706
	(0.404)	(1.144)	(0.00985)	(0.514)	(1.513)	(0.0106)
Machakos	19.57	56.48	0.318	7.417	72.12	0.177
	(1.885)	(2.081)	(0.0521)	(4.195)	(1.974)	(0.0551)
Overall	18.63	44.88	0.150	13.01	66.46	-0.284
	(0.619)	(0.727)	(0.0450)	(1.339)	(0.756)	(0.0681)

Standard errors in parentheses

3.5 Conclusion and implications

Sustainable agricultural productivity growth is essential to broad-based economic progress in agriculture-based economies. This is particularly relevant in many sub-Saharan African countries where agriculture plays a major role as livelihood source for majority of the population. Yet, agricultural productivity growth in these countries is stymied by, among other things, widespread soil infertility mainly because of improper land management practices.

In Kenya, soil infertility combined with agronomic practices that may not effectively respond to fertility needs of the soil is manifest in the stagnation of aggregate maize yield over time despite increasing mineral fertilizer application and planting of improved maize varieties. Because there is virtually no land for expansion to increase agricultural production, the onus is on increasing agricultural productivity through technological change, increased efficiency in use of existing technology and productive resources or both.

This study focused on technical efficiency in maize production in Kenya. The first objective was to understand the level of technical efficiency in smallholder maize farming and its variation across farms, and the factors that are responsible for that variation with a focus on farmers' subjective soil fertility perception. The second was to evaluate the importance of controlling for environmental production conditions (soil conditions and agronomic practices) in agricultural productivity and technical efficiency estimation. The study applied the stochastic production frontier approach due to Meeusen & van Den Broeck (1977) and Aigner et al (1977).

Three key results have emerged from the study. First, maize farmers in Kenya are generally technically inefficient and the level of technical efficiency varies across regions. The mean of technical efficiency is 0.75 when environmental variables are controlled for and 0.70 without controlling for them in the model estimation. On average, the inefficiency-induced foregone maize yield is 0.34 tonnes/acre when environmental variables are controlled for and 0.45 tonnes/acre without controlling for environmental variables. These forgone yield levels are significant considering that the reported (actual) yield averaged only 1.3 tonnes/acre. Secondly, estimation of technical efficiency without controlling for environmental production conditions in the model results in underestimated technical efficiency levels, as evident in the dissimilar results between the model specifications with and without environmental variables. Lastly, results have shown that farmers that view their plots as fertile have, on average, 19% lower technical inefficiency than do those that consider their plots infertile. In addition, farmers whose perceptions about the fertility status of their plots are consistent with measured soil fertility have 2.9% lower technical inefficiency, on average, than do those with perceptions that do not match measured soil fertility.

We conclude that smallholder farmers are technically inefficient and inefficiency varies across regions. Therefore, different regions may require different strategies to improve farmers' management of inputs to raise maize productivity. The result that farmers' perceptions about soil fertility explains variation in technical efficiency underscores the importance of farmers' access to information that can enhance their knowledge and understanding about the correct soil fertility conditions on their farms. Such information would help them make better choices about appropriate inputs and agronomic management practices to apply. Finally, failure to account for environmental production conditions in agricultural productivity analysis may produce results that are less accurate and unreliable.

APPENDIX

Table A3.1: OLS regression results

Response variable: Log of maize yield	With environmental variables		Without environmental variables	
	Coefficient	Std. error	Coefficient	Std. error
Explanatory Variables				
Log of fertilizer (lnN)	0.229	0.160	0.162	0.162
Log of seed (lnseed)	0.639	0.417	0.721*	0.423
Log of labor (lnlabour)	-0.0599	0.186	-0.0645	0.189
lnN x lnN	0.154***	0.0263	0.141***	0.0265
lnseed x lnseed	-0.102	0.200	-0.140	0.203
lnlabour x lnlabour	0.0250	0.0251	0.0281	0.0255
lnN x lnseed	-0.140	0.107	-0.0968	0.109
lnN x lnlabour	-0.0745*	0.0387	-0.0697*	0.0392
lnseed x lnlabour	0.0727	0.125	0.0660	0.127
Improved seed (1=yes)	0.189**	0.0785	0.217***	0.0797
Mechanized land preparation (1=yes)	0.199***	0.0540	0.198***	0.0549
Log of plot size	-0.106***	0.0267	-0.117***	0.0269
No. of crops on plot	0.117***	0.0213	0.109***	0.0182
Year dummy (1=2016)	0.209***	0.0585	0.274***	0.0364
Manure/compost use (1=yes)	0.179***	0.0384		
Maize-legume intercrop (1=yes)	-0.0488	0.0523		
Crops residue use (1=yes)	0.000824	0.0375		
Total organic carbon	0.0277	0.0289		
Phosphorus	-0.00151	0.00107		
Soil pH	0.172***	0.0376		
Sand content	-0.00226	0.00169		
Rainfall	0.0000648*	0.0000390		
Moisture stress	-0.209	0.147		
Transnzoia county dummy[a]	-0.0738	0.0616	-0.0748	0.0588
Kakamega county dummy[a]	-0.304***	0.0758	-0.351***	0.06 30
Kisii county dummy[a]	-0.602***	0.0911	-0.558***	0.0789
Machakos county dummy[a]	-0.275***	0.0878	-0.342***	0.0776
Constant	3.596***	0.791	4.538***	0.763
Adjusted R-squared	0.385		0.359	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.2: Maximum likelihood results of stochastic frontier model with Cobb-Douglas specification

Response variable: Log of maize yield	With environmental variables		Without environmental variables	
	Coef.	Std. error	Coef.	Std. error
Frontier				
Log of fertilizer (lnN)	0.203***	0.0192	0.172***	0.0186
Log of seed (lnseed)	0.470***	0.0570	0.510***	0.0575
Log of labor (lnlabour)	0.0384	0.0234	0.0419*	0.0232
Improved seed (1=yes)	0.106	0.0762	0.147*	0.0786
Mechanized land preparation (1=yes)	0.143***	0.0523	0.133**	0.0532
Log of plot size	-0.156***	0.0298	-0.189***	0.0320
No. of crops on plot	0.110***	0.0209	0.0970***	0.0181
Year dummy (1=2016)	0.209***	0.0553	0.293***	0.0345
Manure/compost use (1=yes)	0.145***	0.0374		
Maize-legume intercrop (1=yes)	-0.0565	0.0506		
Crops residue use (1=yes)	0.0559	0.0360		
Total organic carbon	0.0389	0.0296		
Phosphorus	-0.00164	0.00102		
Soil pH	0.155***	0.0365		
Sand content	-0.00193	0.00164		
Rainfall	0.0000426	0.0000382		
Moisture stress	-0.191	0.138		
Transnzoia county dummy[a]	-0.109*	0.0593	-0.0890	0.0568
Kakamega county dummy[a]	-0.344***	0.0716	-0.380***	0.0604
Kisii county dummy[a]	-0.725***	0.0881	-0.675***	0.0765
Machakos county dummy[a]	-0.372***	0.0852	-0.424***	0.0767
Constant	4.363***	0.314	5.194***	0.204
σ_v^2	-1.624***	0.110	-1.559***	0.123
Inefficiency				
Mean function (μ)				
Plot manager education (yrs)	-0.0471**	0.0234	-0.0462**	0.0183
Plot manager is female (1=yes)	-0.0726	0.182	-0.170	0.123
Soil fertility perception (1=fertile)	-1.430**	0.716	-0.733***	0.185
Farmer perception consistent with soil test (1=yes)	-0.631**	0.280	-0.455***	0.173
Plot size (acres)	-0.146**	0.0659	-0.276***	0.0935
HH landholding (acres)	0.0418*	0.0215	0.0365***	0.0130
Walking time to plot (mins)	-0.00418	0.00802	-0.000623	0.00480
Plot owned (1=yes)	-0.442	0.269	-0.585**	0.260
Distance to extension (km)	0.00498	0.0146	0.0142*	0.00750
Farming experience (yrs)	0.00284	0.00772	-0.00164	0.00551
Constant	1.193***	0.435	1.476***	0.286

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.2 (cont'd)

Response variable: Log of maize yield	With environmental variables		Without environmental variables	
	Coef.	Std. error	Coef.	Std. error
Variance function (σ_u^2)				
Plot manager education (yrs)	-0.0204	0.0300	-0.00747	0.0321
Plot manager is female (1=yes)	-0.103	0.203	0.117	0.214
Soil fertility perception (1=fertile)	0.560*	0.301	-0.0173	0.249
Farmer perception consistent with soil test (1=yes)	0.617***	0.236	0.538*	0.280
Plot size (acres)	0.0570	0.0826	0.167	0.108
HH landholding (acres)	-0.142***	0.0527	-0.157**	0.0798
Walking time to plot (mins)	0.00794	0.00810	0.0115	0.00988
Plot owned (1=yes)	1.133**	0.472	2.666	1.739
Distance to extension (km)	-0.000173	0.0182	-0.0167	0.0178
Farming experience (yrs)	-0.00666	0.00853	-0.00313	0.00828
Constant	-1.700**	0.809	-3.382*	1.970
Log-likelihood	-883.959		-905.144	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.3: Maximum likelihood results of stochastic frontier model of KMHLBC

Response variable: Log of maize yield	With EV		W/o EV	
	Coef.	Std. error	Coef.	Std. error
Frontier				
Log of fertilizer (lnN)	0.228	0.179	0.0819	0.159
Log of seed (lnseed)	0.795**	0.398	0.870**	0.414
Log of labor (lnlabour)	-0.0752	0.184	-0.0424	0.182
lnN x lnN	0.133***	0.0267	0.127***	0.0256
lnseed x lnseed	-0.174	0.192	-0.227	0.200
lnlabour x lnlabour	0.0215	0.0235	0.0130	0.0239
lnN x lnseed	-0.162	0.101	-0.0971	0.107
lnN x lnlabour	-0.0507	0.0459	-0.0282	0.0375
lnseed x lnlabour	0.0740	0.122	0.0590	0.121
Improved seed (1=yes)	0.137	0.0835	0.166**	0.0767
Mechanized land preparation (1=yes)	0.149***	0.0515	0.182***	0.0571
Log of plot size	-0.148**	0.0682	-0.310***	0.0686
No. of crops on plot	0.118***	0.0209	0.115***	0.0178
Year dummy (1=2016)	0.191***	0.0601	0.277***	0.0353
Manure/compost use (1=yes)	0.144***	0.0404		
Maize-legume intercrop (1=yes)	-0.0630	0.0568		
Crops residue use (1=yes)	0.0307	0.0369		
Total organic carbon	0.0302	0.0288		
Phosphorus	-0.00186*	0.00103		
Soil pH	0.164***	0.0364		
Sand content	-0.00191	0.00163		
Rainfall	0.0000419	0.0000380		
Moisture stress	-0.222	0.147		
Transnzoia county dummy[a]	-0.119**	0.0599	-0.0862	0.0563
Kakamega county dummy[a]	-0.325***	0.0734	-0.362***	0.0611
Kisii county dummy[a]	-0.656***	0.0887	-0.644***	0.07 72
Machakos county dummy[a]	-0.346***	0.0995	-0.453***	0.0774
Constant	4.155***	0.947	5.617***	0.761
σ_v^2	-1.730***	0.405	-2.276***	0.363
Inefficiency				
Mean function (μ)				
Plot manager education (yrs)	-0.0628	0.130	-0.0228***	0.00596
Plot manager is female (1=yes)	-0.102	0.318	-0.00623	0.0417
Soil fertility perception (1=fertile)	-0.795	1.929	-0.281***	0.0584
Farmer perception consistent with soil test (1=yes)	-0.221	0.708	-0.0681	0.0531
Plot size (acres)	-0.188	0.316	-0.167***	0.0476
HH landholding (acres)	0.00829	0.0334	-0.00135	0.00556
Walking time to plot (mins)	0.0000886	0.00859	-0.000753	0.00203
Plot owned (1=yes)	0.184	0.319	0.0463	0.0659
Distance to extension (km)	0.00226	0.0115	0.00291	0.00397
Farming experience (yrs)	-0.00474	0.0110	-0.00214	0.00141
Constant	0.884*	0.463	1.600***	0.287
Log-likelihood	-879.972		-901.194	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.4: Maximum likelihood results of various stochastic frontier model of CFCFGH

Response variable: Log of maize yield	With EV		W/o EV	
	Coef.	Std. error	Coef.	Std. error
Frontier				
Log of fertilizer (lnN)	0.233	0.151	0.170	0.152
Log of seed (lnseed)	0.785**	0.393	0.924**	0.400
Log of labor (lnlabour)	-0.0717	0.176	-0.0659	0.179
lnN x lnN	0.132***	0.0251	0.120***	0.0253
lnseed x lnseed	-0.179	0.190	-0.229	0.194
lnlabour x lnlabour	0.0211	0.0230	0.0213	0.0233
lnN x lnseed	-0.154	0.101	-0.125	0.102
lnN x lnlabour	-0.0555	0.0372	-0.0456	0.0377
lnseed x lnlabour	0.0800	0.120	0.0661	0.122
Improved seed (1=yes)	0.150*	0.0767	0.174**	0.0782
Mechanized land preparation (1=yes)	0.151***	0.0513	0.144***	0.0523
Log of plot size	-0.154***	0.0291	-0.168***	0.0291
No. of crops on plot	0.117***	0.0207	0.104***	0.0178
Year dummy (1=2016)	0.189***	0.0557	0.267***	0.0348
Manure/compost use (1=yes)	0.145***	0.0368		
Maize-legume intercrop (1=yes)	-0.0622	0.0501		
Crops residue use (1=yes)	0.0380	0.0356		
Total organic carbon	0.0281	0.0275		
Phosphorus	-0.00186*	0.00101		
Soil pH	0.167***	0.0362		
Sand content	-0.00205	0.00162		
Rainfall	0.0000439	0.0000377		
Moisture stress	-0.226	0.138		
Transnzoia county dummy[a]	-0.116**	0.0585	-0.101*	0.0558
Kakamega county dummy[a]	-0.322***	0.0719	-0.368***	0.0606
Kisii county dummy[a]	-0.660***	0.0880	-0.623***	0.0758
Machakos county dummy[a]	-0.337***	0.0839	-0.404***	0.0754
Constant	4.091***	0.748	4.920***	0.724
σ_v^2	-1.693***	0.117	-1.666***	0.116
Inefficiency				
Variance function (σ_u^2)				
Plot manager education (yrs)	-0.0751***	0.0214	-0.0711***	0.0206
Plot manager is female (1=yes)	-0.177	0.145	-0.113	0.141
Soil fertility perception (1=fertile)	-0.830***	0.220	-0.905***	0.220
Farmer perception consistent with soil test (1=yes)	-0.122	0.179	-0.174	0.179
Plot size (acres)	-0.221**	0.0979	-0.230**	0.0988
HH landholding (acres)	0.000220	0.0169	-0.00153	0.0167
Walking time to plot (mins)	0.00188	0.00708	0.00326	0.00692
Plot owned (1=yes)	0.301	0.234	0.198	0.222
Distance to extension (km)	0.00357	0.0136	0.00799	0.0133
Farming experience (yrs)	-0.00460	0.00469	-0.00521	0.00466
Constant	0.200	0.389	0.364	0.376
Log-likelihood	-878.626		-902.744	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.5: Maximum likelihood results of various stochastic frontier model of Stevenson (1980)

Response variable: Log of maize yield	With EV		W/o EV	
	Coef.	Std. error	Coef.	Std. error
Frontier				
Log of fertilizer (lnN)	0.167	0.156	0.0984	0.158
Log of seed (lnseed)	0.834**	0.406	1.004**	0.410
Log of labor (lnlabour)	-0.110	0.183	-0.104	0.184
lnN x lnN	0.147***	0.0256	0.134***	0.0258
lnseed x lnseed	-0.186	0.193	-0.239	0.196
lnlabour x lnlabour	0.0261	0.0242	0.0302	0.0244
lnN x lnseed	-0.154	0.102	-0.118	0.104
lnN x lnlabour	-0.0433	0.0394	-0.0352	0.0404
lnseed x lnlabour	0.0728	0.124	0.0477	0.126
Improved seed (1=yes)	0.160**	0.0770	0.183**	0.0790
Mechanized land preparation (1=yes)	0.179***	0.0516	0.168***	0.0525
Log of plot size	-0.0981***	0.0256	-0.107***	0.0257
No. of crops on plot	0.118***	0.0208	0.104***	0.0179
Year dummy (1=2016)	0.190***	0.0568	0.255***	0.0351
Manure/compost use (1=yes)	0.166***	0.0375		
Maize-legume intercrop (1=yes)	-0.0696	0.0513		
Crops residue use (1=yes)	0.0205	0.0365		
Total organic carbon	0.0385	0.0282		
Phosphorus	-0.00147	0.00105		
Soil pH	0.165***	0.0367		
Sand content	-0.00159	0.00165		
Rainfall	0.0000504	0.0000381		
Moisture stress	-0.224	0.142		
Transnzoia county dummy[a]	-0.0962	0.0599	-0.0916	0.0570
Kakamega county dummy[a]	-0.307***	0.0733	-0.362***	0.0616
Kisii county dummy[a]	-0.586***	0.0881	-0.568***	0.0762
Machakos county dummy[a]	-0.265***	0.0843	-0.332***	0.0755
Constant	4.167***	0.781	4.963***	0.754
σ_v^2	-1.728***	0.136	-1.649***	0.142
Log-likelihood	-912.054		-937.162	
Observations	1102		1102	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

REFERENCES

REFERENCES

Abdulai, A., & Tietje, H. (2007). Estimating technical efficiency under unobserved heterogeneity with stochastic frontier models: application to northern German dairy farms. *European Review of Agricultural Economics*, *34*(3), 393–416.

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

Ali, M., & Byerlee, D. (1991). Economic efficiency of small farmers in a changing world: A survey of recent evidence. *Journal of International Development*, *3*(1), 1–27.

Alvarez, A., Amsler, C., Orea, L., & Schmidt, P. (2006). Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics, 201–212. https://doi.org/10.1007/s11123-006-7639-3

Ariga, J., & Jayne, T. S. (2009). *Private Sector Responses to Public Investments and Policy Reforms The Case of Fertilizer and Maize Market Development in Kenya* (IFPRI Discussion Paper No. 921).

Barrett, C. B. (1996). On price risk and the inverse farm size – productivity relationship. *Journal of Development Economics*, *51*, 193–215. https://doi.org/10.1016/S0304-3878(96)00412-9

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity-size relationship. *World Development*, *38*(1), 88–97. https://doi.org/10.1016/j.worlddev.2009.06.002

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, *38*(3), 387–399. https://doi.org/10.1016/0304-4076(88)90053-X

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

Belotti, F., & Ilardi, G. (2015). Consistent inference in fixed-effects stochastic frontier models.

Benjamin, D. (1995). Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics*, *46*(1), 51–84. https://doi.org/10.1016/0304-3878(94)00048-H

Besley, T. (1995). Property Rights and Investment Incentives: Theory and Evidence from Ghana. *Journal of Political Economy*, *103*(5), 903–937. https://doi.org/10.1086/262008

Bot, A., & Benites, J. (2005). *The importance of soil organic matter: key to drought-resistant soil and sustained food production*. *FAO Soils Bulletin*. https://doi.org/10.1080/03650340214162

Carletto, C., Savastano, S., & Zezza, A. (2013). Fact or artifact: The impact of measurement errors on the farm size-productivity relationship. *Journal of Development Economics*, *103*, 254–261. https://doi.org/10.1016/j.jdeveco.2013.03.004

Caudill, S. B., & Ford, J. M. (1993). Biases in frontier estimation due to heteroscedasticity. *Economics Letters*, *41*(1), 17–20.

Caudill, S. B., Ford, J. M., & Gropper, D. M. (1995). Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroscedasticity. *Journal of Business & Economic Statistics*, *13*(1), 105–111. https://doi.org/10.1080/07350015.1995.10524583

Chamberlain, G. (1984). Panel Data. In *Handbook of Econometrics* (Vol. II, pp. 1247–1318). https://doi.org/10.1016/S1573-4412(84)02014-6

Chen, Y., Schmidt, P., & Wang, H. (2014). Consistent estimation of the fixed effects stochastic frontier model. *Journal of Econometrics*, *181*(2), 65–76. https://doi.org/10.1016/j.jeconom.2013.05.009

Coelli, T. J. (1995). Recent developments in frontier modelling and efficiency measurement. *Australian Journal of Agricultural Economics*, *39*(3), 219–245.

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

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

Diao, X., Hazell, P., Resnick, D., & Thurlow, J. (2007). *The Role of agriculture in development: Implications for sub-Saharan Africa*. Washington, D.C., USA. https://doi.org/153

Doran, J. W., & Parkin, T. B. (1994). Defining and assessing soil quality. *Defining Soil Quality for a Sustainable Environment*.

Ekbom, A., Alem, Y., & Sterner, T. (2013). Integrating soil science into agricultural production frontiers. *Environment and Development Economics*, *18*, 1–18. Retrieved from http://journals.cambridge.org/abstract_S1355770X13000089

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

Farsi, M., Filippini, M., & Kuenzle, M. (2005). Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. *Applied Economics*, *37*, 2127–2141.

Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency : a maximum simulated likelihood approach. *Journal of Productivity Analysis*, *45*, 187–196.

Filippini, M., Hunt, L. C., & Zorić, J. (2014). Impact of energy policy instruments on the estimated level of underlying energy ef fi ciency in the EU residential sector. *Energy Policy*, *69*, 73–81.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., … Michaelsen, J. (2015). Climate Harzads Group. https://doi.org/10.1038/sdata.2015.66

Government of Kenya. (2014). *Soil suitability evaluation for maize production in Kenya soil suitability evaluation for*. Nairobi, Kenya.

Greene, W. H. (1980). Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics*, *13*, 27–56.

Greene, W. H. (2005). Fixed and random effects in stochastic frontier. *Journal of Productivity Analysis*, *23*, 7–32.

Griffiths, W. E., & Hajargasht, G. (2016). Some models for stochastic frontiers with endogeneity. *Journal of Econometrics*, *190*, 341–348.

Grisso, R., Alley, M., Holshouser, D., & Thomason, W. (2009). Precision farming tools: soil electrical conductivity. *Virginia Cooperative Extension*, *442*(508), 1–6. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Precision+farming+tools:+soil+electrical+conductivity#7

Hadri, K. (1999). Estimation of a Doubly Heteroscedastic Stochastic Frontier Cost Function. *Journal of Business & Economic Statistics*, *17*(3), 359–363. https://doi.org/10.1080/07350015.1999.10524824

Haggblade, S., & Hazell, P. B. R. (Eds.). (2010). *Successes in African agriculture: Lessons for the future*. Baltimore: The Johns Hopkins University Press.

Henningsen, A. (2014). Introduction to econometric production analysis with R (Draft version). Department of Food and resource Economics, University of Copenhagen.

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

Jondrow, J., Lovell, C. A. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical efficiency in the stochastic frontier production fucntion model. *Journal of Econometrics*, *19*, 233–238.

Kagin, J., Taylor, J. E., & Yúnez-Naude, A. (2015). Inverse Productivity or Inverse Efficiency? Evidence from Mexico. *The Journal of Development Studies*, *388*(December), 1–16. https://doi.org/10.1080/00220388.2015.1041515

Kagin, J., Taylor, J. E., & Yúnez-Naude, A. (2016). Inverse Productivity or Inverse Efficiency? Evidence from Mexico. *The Journal of Development Studies*, *52*(3), 396–411. https://doi.org/10.1080/00220388.2015.1041515

Kibaara, B. W. (2005). *Technical efficiency in Kenya's maize production: An application of the stochastic frontier approach*. Colorado State University.

Kumbhakar, C. S., & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

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

Kumbhakar, S. C., Ghosh, S., & McGuckin, J. T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in US . dairy farms. *Journal of Business & Economic Statistics*, *9*(3), 279–286.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models : a study of Norwegian grain farming. *Journal of Productivity Analysis*, *41*, 321–337.

Kumbhakar, S. C., Wang, H.-J., & Horncastle, A. P. (2015). *A practitioner's guide to stochastic frontier analysis using stata*. New York: Cambridge University Press.

Kumbhakar, S. C., Parmeter, C. F., & Tsionas, E. G. (2013). A zero inefficiency stochastic frontier model. *Journal of Econometrics*, *172*(1), 66–76.

Lal, R. (2006). Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, *209*(August 2005), 197–209.

Liu, Y., & Myers, R. (2009). Model selection in stochastic frontier analysis with an application to maize production in Kenya. *Journal of Productivity Analysis*, *31*, 33–46. https://doi.org/10.1007/s11123-008-0111-9

Maguire, R., Mark, A., & Flowers, W. (2009). *Fertilizer Types and Calculating Application Rates* (Publication No. 424–35). *Virginia Cooperative Extension*. Communications and Marketing, College of and Life Sciences, Virginia Polytechnic Institute and State University.

Marenya, P. P., & Barrett, C. B. (2007). Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. *Food Policy*, *32*(4), 515–536.

Marenya, P. P., & Barrett, C. B. (2009). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, *40*, 561–572.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error, *18*(2), 435–444.

Mghenyi, E. (2015). *The impact of agricultural credit on demand for factors of production , farm output , and profitability in Kenya*. Michigan State University.

Mochebelele, M. T., & Winter-Nelson, A. (2000). Migrant labor and farm technical efficiency in Lesotho. *World Development*, *28*(1), 143–153.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, *46*(1), 69. https://doi.org/10.2307/1913646

Parmeter, C. F., & Kumbhakar, S. C. (2014). Efficiency analysis : A Primer on recent advances. *Foundations and Trends in Econometrics*, *7*(3–4), 191–385.

Pitt, M. M., & Lee, L.-F. (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*, *9*, 43–64.

Reifschneider, D., & Stevenson, R. (1991). Systematic departures from the frontier : A framework for the analysis of firm inefficiency. *International Economic Review*, *32*(3), 715–723.

Rho, S., & Schmidt, P. (2015). Are all firms inefficient? *Journal of Productivity Analysis*, *43*(3), 327–349.

Sanchez, P. A., Shepherd, K. D., Soule, M. J., Place, F. M., Buresh, R. J., Izac, A. N., & Kwesiga, F. R. (1997). Soil Fertility Replenishment in Africa : An Investment in Natural Resource Capital. *American Society of Agronomy and Soil Science Society of America*, (51), 1–46.

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

Sheahan, M. B. (2011). *Analysis of fertilizer profitability and use in Kenya*. Michigan State University.

Sherlund, S. M., Barrett, C. B., & Adesina, A. A. (2002). Smallholder technical efficiency controlling for environmental production conditions, *69*, 85–101.

Smale, M., & Olwande, J. (2014). Demand for maize hybrids and hybrid change on smallholder farms in Kenya. *Agricultural Economics (United Kingdom)*, *45*(4).

Smaling, E. M. A., Nandwa, S. M., & Janssen, B. H. (1997). Soil Fertility in Africa Is at Stake. *American Society of Agronomy and Soil Science Society of America*, (51).

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

Stoorvogel, J. J., & Windmeijer, P. N. (1993). Calculating soil nutrient balances in Africa at different scales. *Fertilizer Research*, *35*, 237–250.

Stoorvogel, J., & Smaling, E. M. A. (1998). Research on soil fertility decline in tropical environments : integration of spatial scales. *Nutrient Cycling in Agroecosystems 50:*, *50*, 151–158.

The Montepellier Panel. (2013). *Sustainable Intensification: A New Paradigm for African Agriculture*. London.

The World Bank. (2007). *World development report 2008: Agriculture for development*. Washington, DC: The World Bank.

Tiedemann, T., & Latacz-Lohmann, U. (2013). Production Risk and Technical Efficiency in Organic and Conventional Agriculture – The Case of Arable Farms in Germany. *Journal of Agricultural Economics*, *64*(1), 73–96.

Tittonell, P., Corbeels, M., Wijk, M. T. Van, Vanlauwe, B., & Giller, K. E. (2008). Combining Organic and Mineral Fertilizers for Integrated Soil Fertility Management in Smallholder Farming Systems of Kenya: Explorations Using the Crop-Soil Model FIELD. *Agronomy Journal*, *100*(5), 1511–1526.

Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, *143*, 76–90. https://doi.org/10.1016/j.fcr.2012.10.007

Tran, K. C., & Tsionas, M. G. (2015). Zero-inefficiency stochastic frontier models with varying mixing proportion: A semiparametric approach. *European Journal of Operational Research*.

Wang, H. (2003). A Stochastic Frontier Analysis of Financing Constraints on Investment A Stochastic Frontier Analysis of Financing Constraints on Investment : The Case of Financial Liberalization in Taiwan, *21*(3), 406–419. https://doi.org/10.1198/073500103288619016

Wang, H.-J. (2002). Heteroscedasticity and Non-Monotonic Efficiency Effects of a Stochastic Frontier Model. *Journal of Productivity Analysis*, *18*, 241–253.

Wang, H.-J., & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, *18*, 129–144.

Wang, H., & Ho, C. (2010). Estimating fixed-effect panel stochastic frontier models by. *Journal of Econometrics*, *157*(2), 286–296. https://doi.org/10.1016/j.jeconom.2009.12.006

Willy, D. K., Muyanga, M., & Jayne, T. (2016). Can economic and environmental benefits associated with agricultural intensification be sustained at high population densities ? A farm level empirical analysis. In *Invited paper presented at the 5th International Conference of the African Association of Agricultural Economists, September 23-26, 2016, Addis Ababa, Ethiopia*.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data. *MIT Press*. https://doi.org/10.1515/humr.2003.021

Yang, J., Wang, H., Jin, S., Chen, K., Riedinger, J., & Peng, C. (2016). Migration, local off-farm employment, and agricultural production efficiency: evidence from China. *Journal of Productivity Analysis*, *45*(3), 247–259. https://doi.org/10.1007/s11123-015-0464-9