ANALYSIS OF FERTILIZER PROFITABILITY AND USE IN KENYA

By

Megan Britney Sheahan

A THESIS

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

MASTER OF SCIENCE
Agricultural, Food and Resource Economics

2011

ABSTRACT

ANALYSIS OF FERTILIZER
PROFITABILITY AND USE IN KENYA

By

Megan Britney Sheahan

Despite upward trends in fertilizer use on maize fields in Kenya over the past twenty years, it is still widely viewed that fertilizer use is not expanding quickly enough and that application rates are not high enough to meet national food security and agricultural development goals. This thesis takes a critical look at the profitability and use of fertilizer with respect to maize in Kenya using five waves of household level panel data across thirteen years. I estimate a maize yield response model at the field level to ascertain district and soil group level fertilizer response rates by year, then use these estimates to calculate marginal and average value cost ratios under a number of household specific relative price scenarios including consideration of the transport cost of fertilizer and both the buying and selling prices of maize. I compare these profitability metrics and calculated optimal fertilizer application rates to actual fertilizer use values to learn that households in the highest potential areas are using fertilizer at or beyond the most profitable levels while households in the more marginal lowlands areas have steadily approached optimal use levels, with a small gap remaining in 2010. While fertilizer use could be expanded in the lowlands areas, lower application rates might be the most profitable strategy in other areas. When limiting my sample to only areas where fertilizer use is profitable, I estimate a probit model to determine the factors associated with not using commercial fertilizer on maize fields. I find that long distances to the nearest fertilizer seller and relatively adverse nitrogen to maize price ratios are the major deterrents to fertilizer use where otherwise profitable.

ACKNOWLEDGEMENTS

This thesis is truly the collective work of many individuals. I would like to thank my major advisor, Dr. Thom Jayne, for his guidance throughout my master's program and for allowing the opportunity to be involved with this important work in Kenya. I owe tremendous thanks to Dr. Roy Black, without which this analysis would not be nearly as rich or complete. Your constant feedback was invaluable to this process. I also benefited a great deal from the guidance of Dr. Sieglinde Snapp. I appreciate your patience and consistency in helping me to understand the importance of soil. To all three committee members, I thank you for your flexibility and willingness in allowing me to work remotely on this project.

I would also like to thank Joshua Ariga, not only for providing a solid platform off of which to build, but also for being so willing to help me throughout my research. Furthermore, I appreciate all of the guidance and context provided to me by my Kenyan colleagues in AFRE: Joshua Ariga, Milu Muyanga, Miltone Ayieko. Your perspectives and eagerness to answer my many questions are much appreciated. I would also like to thank my colleagues at Tegemeo, specifically John Olwande and Francis Karin, for providing help and advice whenever I asked. It was a great pleasure to participate in data collection and, specifically, to learn from Joseph Opiyo and our team of enumerators. Those several weeks in western Kenya taught me more about agriculture and the challenges of quality data collection than I ever thought possible.

I would also like to thank a number of other individuals who made this thesis better: Margaret Beaver for your thoroughness with this data and your willingness to answer any data management question; Tavneet Suri for the countless hours you put in to bring us quality rainfall data; David Mather for help thinking through and calculating rainfall expectations; Mary

Ollenburger for a wonderful orientation on how to approach soil classification; Dr. Jeff Wooldridge for your advice on estimation and sample selection; Tom Awuor and Dr. Val Kelly for your help gathering information on the fertilizer subsidy programs; Nicky Mason and Jake Ricker-Gilbert for answering my estimation and Stata questions time and time again; and Dr. Swinton, Dr. Loveridge, Luke Reese and Linda Beck for helping to coordinate my remote defense.

And, finally, I would like to thank Graham not only for allowing me to take over dinner time conversation with stories of fertilizer response in Kenya, but also for allowing me to take over the dinner table, quite literally, with all things related to this thesis. Your turn.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1. Motivation

In the past several years, the promotion of fertilizer has become a resounding theme across SSA, particularly following the first African Fertilizer Summit in Abuja, Nigeria in mid-2006. A resurgence of interest in fertilizer use has led to the renewal of large-scale fertilizer subsidy programs across a growing number of countries—Malawi, Nigeria, Zambia, Tanzania, Ghana—and a refocusing on agricultural input intensification by major donors and development programs. Despite increased rhetoric surrounding fertilizer use, serious attention has not been paid to understanding the correlation between the profitability of fertilizer application and observed use patterns. Without a keen understanding of where fertilizer use is actually profitable, fertilizer subsidy and development programs aimed at encouraging fertilizer use are unlikely to stimulate agricultural productivity in a manner congruent with expectations. Only with disaggregated estimates of profitability can one begin to investigate whether or not a real gap exists between where it is profitable for farmers to use fertilizer and where we observe them using it and, if such a gap does exist, what the reasons or constraints to fertilizer use might be.

This thesis determines optimal fertilizer use rates on maize fields and then assesses the degree to which these optimal use rates compare with farmers' actual fertilizer use rates. In Kenya, agricultural and food market liberalization in the mid-1990s contributed to massive new private investment in fertilizer retailing in rural Kenya and a substantial decrease in the prices of both maize and fertilizer, all of which was achieved largely without government subsidies. Largely as a consequence of lower distances from the farm to private fertilizer retailers and lower real fertilizer prices over time, national fertilizer consumption doubled between 1990/91 and 2007/08 (Ministry of Agriculture 2008) with growth not only driven by large-scale farmers but also small-holder farmers (Ariga et al. 2006; 2008; Ariga and Jayne 2009). Using nationwide farm panel data from Kenya, I find that commercial fertilizer was used on about 90 percent of maize fields across most high potential maize areas of western and central Kenya in 2010. The percentage of fields fertilized in all eastern and western lowlands areas has increased from about 11 percent to 40 percent between 1997 and 2010, with tremendous variation across districts. Despite upward trends in fertilizer use on maize fields in Kenya over the past twenty years, the Government of Kenya has contended that fertilizer use is not expanding quickly enough and that application rates are not high enough, as evidence from the creation of a comprehensive multi-million dollar fertilizer and improved seed subsidy and training program, the National Accelerated Agricultural Inputs Access Program (NAAIAP). Before further policy emphasis is placed on increasing fertilizer use, analysis is needed on how actual use patterns compare with calculated profitability levels and to identify if and where a legitimate gap remains between the two.

2. Literature Gap and Objectives

Fertilizer use in Kenya is a well-studied topic. Duflo et al. (2008) use randomized on-farm trials in Busia district of Western Kenya and the total increase in revenue from fertilizer indexed to the price of the input to show that fertilizer use is profitable at a range of different application levels, although observe few farmers in the area actually using it. Marenya and Barrett (2009b), too, focus on small farms in Western Kenya but with specific interest in how the initial soil organic matter composition of a particular plot relates to fertilizer response and yields, finding that insufficient available soil organic matter likely limits the usefulness of applied inorganic fertilizer. On average, they find fertilizer use to be profitable on plots in their sample area, defined as where the marginal value product of fertilizer exceeded its market price. Moreover, they find an average nitrogen application rate of only 5.2 kilograms per hectares with 88 percent of farmers applying some fertilizer. Matsumoto and Yamano (2011) use two waves of panel data from mostly western and central Kenya to look at fertilizer profitability with similar interest in soil quality at the plot level, and find that farmers in Kenya used fertilizer at estimated economically optimal levels in one of the two survey years. Using experimental data from 70 sites in the late-1980s, Hassan et al. (1998) study fertilizer profitability under pre-liberalization market conditions and prices. Other studies focus on the institutional and behavioral elements of fertilizer use. Alene et al. (2008) investigate the role of transaction costs in suppressing fertilizer use, noting that the increasing cost of information limits farmers' access to fertilizer. Duflo et al. (2009) show how farmers in Western Kenya are prone to behavioral biases, namely procrastination, limiting an otherwise profitable fertilizer use decision.

While useful in conceptualizing the fertilizer profitability and use decisions of farmers in Kenya, these studies focus on very limited geographic areas, derive their estimates using data collected over relatively short periods of time, or require updating to account for current market conditions. Furthermore, these studies confine their analyses to areas of western or central Kenya where fertilizer use is already high, forgoing analysis on the eastern part of the country where the number of users has increased steadily over the past several years. No study, to my knowledge, utilizes a long time series over which profitability conditions have likely shifted in order to study profitability conditions over many years and across all maize producing areas in Kenya. This thesis, then, will investigate fertilizer use and profitability across Kenya using variation over time (5 waves of panel data covering a 13 year time span) and space (120 villages in over 24 districts) covering a large number of maize producing areas. In doing so, I ask the following:

- How does the response of maize to fertilizer application vary across Kenya? What are the impacts of specific field-level, household, community, and agro-ecological factors on maize response and maize response to fertilizer use?
- Are households in Kenya using fertilizer on maize fields where it is profitable to do so? Is there room for profitably expanding fertilizer use in certain areas? How have changes in relative prices affected where fertilizer use is profitable? How does incorporating the transportation cost of fertilizer affect its profitability? How does the maize marketing position of a household (i.e., net buyer or net seller) affect fertilizer profitability?
- What are economically optimal levels of fertilizer application? For those households that are using fertilizer on maize fields, are they doing so at these economically optimal levels? Or, does a gap exist between optimal and observed fertilizer application rates?
- What are the characteristics of households not using commercial fertilizer on maize where it is profitable? Do these characteristics mimic the constraints to input use often described in the input adoption literature (e.g., credit and information constraints)?

Using a nationally representative household panel dataset, I estimate fertilizer profitability then compare with observed fertilizer use patterns over time. I look at a number of different profitability scenarios to see how changes in input and output prices, transportation costs and farmers' position in the maize market (e.g., net buyer versus net seller of maize) affect fertilizer profitability. While cognizant of the loss of household-level specificity, I estimate district level optimal fertilizer use rates and compare with actual use levels to establish where a gap exists between observed and estimated economically optimal levels. Using data from over a seven year period, I also look at how fertilizer profitability and use decisions are affected by household- and village-specific attributes using reasons provided by households and a binary response model where the sample is limited to only those places where fertilizer use is estimated to be profitable.

3. Existing Literature on Fertilizer Use and Trends

Before specifying my conceptual framework and methodology, I briefly review the existing literature on the topic. Here, I discuss why inorganic fertilizer is the input of focus, detail what profitability and use analysis has been done across sub-Saharan Africa, and describe national fertilizer use trends in Kenya.

3.1. Why inorganic fertilizer?

The primary aim of applying inorganic fertilizer is to increase the biological base of the plant system (Weight and Kelly 1999). In doing so, inorganic fertilizer affords both plant productivity gains and longer-term replenishment of nutrients back into the soil. While the former reason is of main interest to this analysis, research shows that the two are also inextricably linked. In Ethiopia, Yesuf et al. (2005) find that land degradation due to soil fertility depletion can cause significant decreases in agricultural productivity. With evidence from Western Kenya, Marenya and Barrett (2009b) show that fertilizer profitability is contingent upon soil fertility levels, meaning farmers with poor soils are less likely to use fertilizer and get caught in the “trap” of low productivity due to the quality of their soil (i.e., soil structure, pore space, water-holding capacity, ability to release nutrients into the soil).

With respect to agricultural productivity, Morris et al. (2007) find that low agricultural growth in Africa is positively correlated with and explained in large part by low fertilizer use. Furthermore, a number of studies show the importance of fertilizer use in agricultural productivity gains in other parts of the world. Research shows, for instance, that over 50 percent of the productivity gains experienced in Asia during the Green Revolution can be attributed to increased fertilizer use, not just improved seed (Hopper 1993; Tomich et al. 1995). It is noted that water control afforded by irrigation was a major contributor to fertilizer's contribution to productivity growth in Asia (Gulati and Narayanan 2003; Johnson et al. 2003). Worldwide, Bumb (1995) finds that one-third of growth in cereal production can be attributed to fertilizer use. Overall, the contribution of inorganic fertilizer to yields and, subsequently, increased agriculturally productivity is not disputed.

The depletion of nutrients from the soil is also a major issue. Only about 20 percent of the land in Kenya is considered medium to high potential agricultural land (Tabu et al. 2007). With high population growth, particularly in the agriculturally productive areas, farmers are forced not only to cultivate suboptimal agricultural land (the other 80 percent of land), but also to use the same plots of land season after season without replenishing the soils through fallowing. Drechsel et al. (2001) analytically show the strong significant relationship between population pressure, reduced fallow periods and soil nutrient depletion, much like what is happening in Kenya. Across all of SSA, Stoorvogel and Smaling (1990) estimate that an average of 660 kilograms of nitrogen per hectare, 75 kilograms of phosphorous per hectare, and 450 kg of potassium per hectare have been lost since the 1960s from about 200 million hectares of cultivated land. Similar trends are observed in Kenya.

Traditional African coping strategies (e.g., fallowing, opening new lands, intercropping, mixed crop-livestock) are not capable of adjusting quickly enough to rapid population growth combined with decreasing farm size and decreasing soil fertility (Cleaver and Schreiber 1994). Putting nutrients back into the soil, then, is the only realistic way to maintain the soil health necessary for sustained agricultural production. Fertilizer use is considered the obvious way to overcome soil fertility depletion given high levels of nitrogen and phosphorous content. Similarly, fertilizer itself helps to sustain soil fertility by maintaining particular nutrient pools and by generating additional biomass that is returned to the soil (e.g., via crop residue incorporation, mulching, composting, or manure from livestock grazing harvested fields), thereby sustaining and possibly increasing soil organic matter (Weight and Kelly 1999). Traditional organic fertilizers (i.e., manure and compost) can be used to fix nutrients back into the soil (i.e., plants do not discriminate between organic and inorganic nutrient ions) but a much larger volume is required to do so. For example, most animal manure and plant material contain between 1 and 4 percent nitrogen content compared with 20 to 46 percent in inorganic fertilizers, and the phosphorous content of plant residuals and manure are generally not sufficient to meet crop growth requirements (Sanchez et al. 1997). Morris et al. (2007) claim that simply not enough organic fertilizer exists to "fix" soil nutrient problems in Africa. Nitrogen in organic fertilizer also mineralizes more slowly than inorganic fertilizer, meaning not necessarily consistent with crop growth cycles (Byrnes 1990). Organic fertilizer use is generally recommended in addition to, not as a replacement for, inorganic fertilizer (Weight and Kelly 1999).

Inorganic fertilizer has come to acquire a bad name in more developed countries due to its publicized harmful effects on the environment and health including nitrogen leaching, ammonia volatilization (related to acid rain), the emission of nitrous oxides, and the eutrophication of aquatic environments resulting from phosphorous run-off (see Shaviv and Mikkelsen 1993 for a review). These unfortunate consequences, however, are linked primarily to overuse of fertilizer, not fertilizer use in general (Byrnes 1990; Sanchez et al. 1997), making it that much more imperative to derive and disseminate well-approximated optimal fertilizer application rates. Instead, the main environmental concerns in sub-Saharan Africa currently stem from the rapid depletion of nutrients from the soil, of which fertilizer application is a viable, if incomplete, prescription (Larson and Frisvold 1996). For example, soil fertility depletion leads to increased soil erosion and, thereafter, unwanted sedimentation, siltation of coastal areas and eutrophication of rivers and lakes (Sanchez et al. 1997).

3.2. Fertilizer use and profitability in sub-Saharan Africa

Fertilizer application rates in SSA are far below any other region in the world. Minot and Benson (2009) find that the average fertilizer application rate was only 13 kg/ha in 2008, compared with an average 94 kg/ha in other developing countries. While operating and biophysical environments are considerably different between places, this statistic has prompted a considerable discussion about low fertilizer use in SSA. Researchers provide a long list of reasons why this might be the case. Several articles divide potential reasons for low fertilizer use into demand and supply side factors (Crawford et al. 2003; Morris et al. 2007). On the demand side, both perceived profitability and ability to pay are thought to contribute to low use. Profitability could be hindered by variability in prices (of fertilizer and output) and yield, agro-ecological conditions (i.e., soil characteristics and weather patterns), and lack of knowledge about how properly to use fertilizer. Ability to pay reflects both low income levels and lack of access to credit in many rural areas. On the supply side, having fertilizer available in appropriately sized packaged at the necessary time of year often prohibits access at the farm level (Larson and Frisvold 1996). Kherallah et al. (2002) add that fertilizer costs are higher in Africa than other regions due mostly to high transport costs making it more difficult for poor farmers to obtain. Similarly, they state that Africa does not have the irrigation infrastructure of many other regions which hinders the ability for plants to uptake nutrients in a timely manner. Also, population density is much lower than other places requiring less need for land-saving technologies. Most of these reasons, both on the demand and supply sides, have underlying structural determinants and often can be overcome with appropriate public sector interventions.

In their review, Morris et al. (2007) find fertilizer use to be unprofitable in many parts of Africa due to high prices and transportation costs. Heisey and Mwangi (1997) showed that profitability of fertilizer application to maize, calculated as a ratio of fertilizer price to maize market price, had increased over time in many major maize producing countries in Africa. Meertens (2005) calculated profitability using another metric, value cost ratios (VCR), and found a similar downward trend in profitability, reaching critically low levels particularly in SSA. Yanggen et al. (1998) find that while overall agronomic response to fertilizer in many parts of Africa is similar to other places in the world, the ratio of fertilizer price to output price is much higher, making it one of the least profitable places to purchase the input. Clearly, then, the price at which fertilizer can be procured is an essential component to its profitability and likely use. In a review of four countries in SSA from 1971 to 2001, Heisey and Norton (2007) find that the price of nitrogen was below the world average price at the beginning part of the period but much higher towards the end. This finding is consistent with other claims of falling profitability over time.

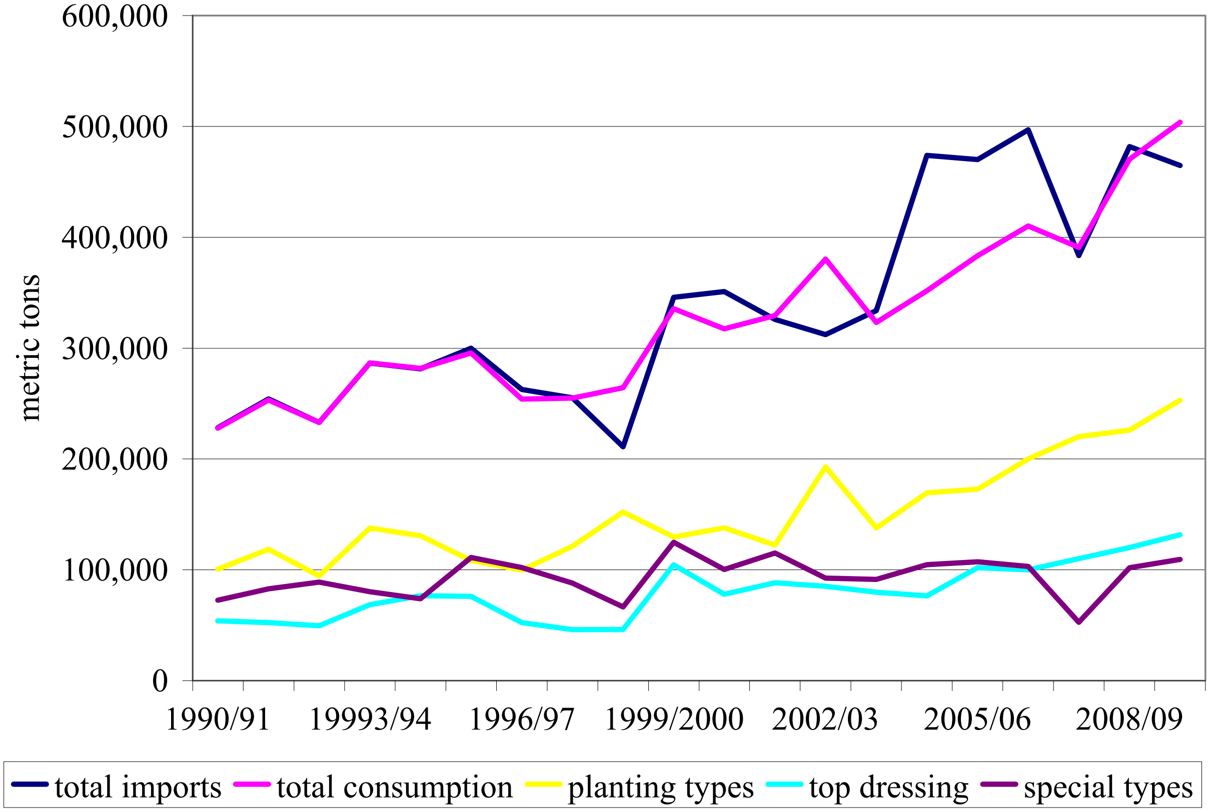
3.3. Fertilizer trends in Kenya

Aggregate trends of SSA may be unimpressive, but country level statistics show greater variation and some success stories, Kenya among them. Ariga et al. (2006) group countries in

Africa by intensity of fertilizer use and percentage growth in fertilizer amount and find that of the four countries which use an average of 25 kilograms per hectare, three have had a growth rate of less than 30 percent over the 1990-2003 period (Swaziland, Malawi, and Zimbabwe) while one (Kenya) has had both high use and high growth. Ariga et al. (2008), using a nationally representative panel, find the percentage of smallholder farmers using fertilizer on maize to have increased from 56 percent in 1996 to 70 percent in 2007 coupled with an increase in application amount from 34 kilograms per acre in 1996 to 45 kilograms per acre in 2007, with statistically significant variation across regions and districts, as expected. For example, in Nakuru district, a high potential maize area, Obare et al. (2003) found over 90 percent of farmers using fertilizer on maize. To the west in Vihiga and South Nandi Districts, Marenya and Barrett (2009) found that 88 percent of the 260 farmers sampled in their study used fertilizer in the 2004 main crop season. In lower potential and semi-arid areas, like the Coastal and Western Lowlands, Ariga et al. (2008) find fertilizer use still below 15 percent, likely as a result of a very different response and market environment.

Figure 1 summarizes national-level trends over time. Notice that between the mid-1990s and 2005, fertilizer consumption increased by about one-third. Then from 2005 to 2010, fertilizer consumption again increased by one-fourth. The momentary drop in both fertilizer consumption and imports in the 2007/08 season is attributed to both high international prices and the post-election violence in Kenya. Ariga et al. (2006) explain which conditions in Kenya have been the sources of impressive growth in fertilizer use including a stable fertilizer policy environment, a reduction in marketing margins following liberalization, a major increase in the number of fertilizer retailers operating in rural areas (reducing the average distance traveled from farm to acquisition source), and a noticeable shift from monocropping to intercropping in some areas.

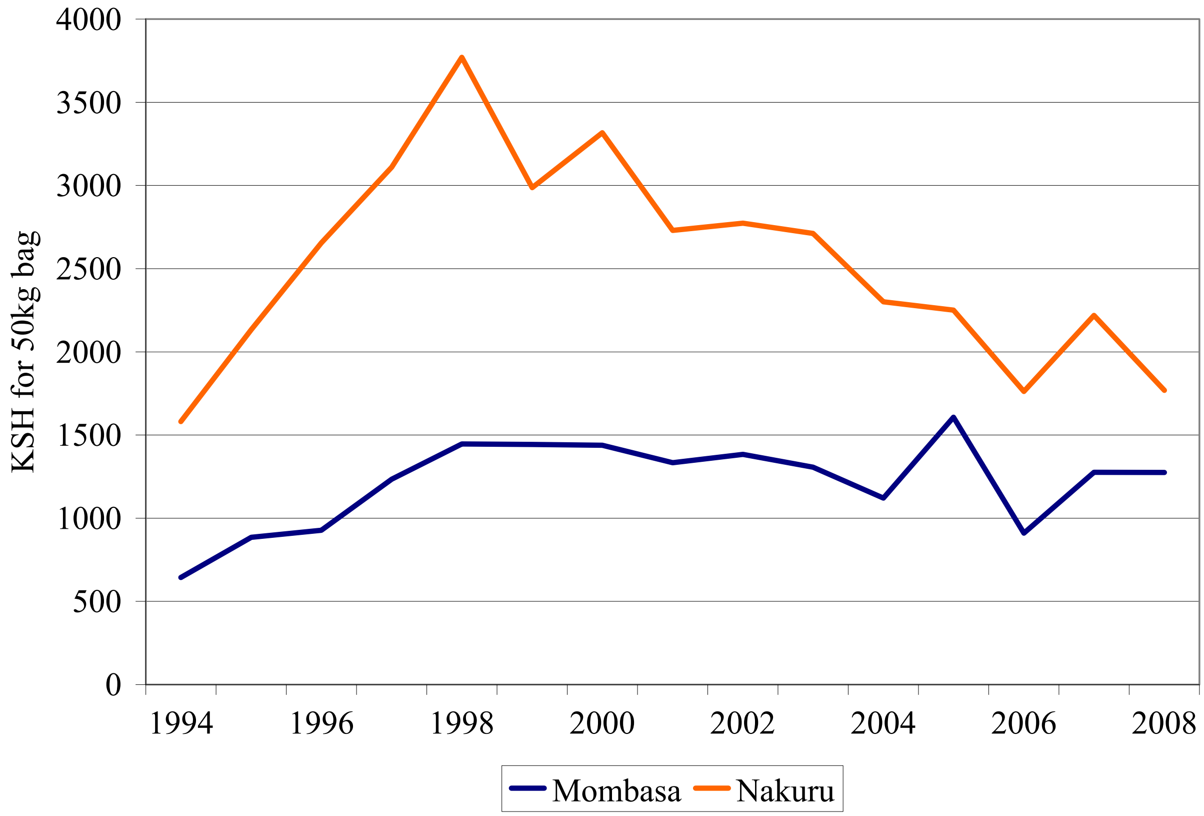
Figure 1: National level fertilizer consumption and imports over time



Source: Ministry of Agriculture in Kenya.
Note: For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

Like many other African countries, virtually all fertilizer consumed in Kenya is imported (see Figure 1). This makes fertilizer prices particularly susceptible to swings in international commodity prices. Imported fertilizer arrives at the port in Mombasa and makes its way to the more agriculturally productive areas in central and western Kenya via private traders and the government. Figure 2 shows the trends in price of fertilizer observed at Mombasa and Nakuru; the difference between the two represents the margins absorbed by traders, transporters, packagers and marketers. In general, prices in Mombasa (representing international prices plus port charges) has stayed constant over time while prices in Nakuru has fallen dramatically since the late 1990s, signaling a reduction in fertilizer marketing margins over time. By asking key informants in the fertilizer sector, Ariga et al. (2008) report four reasons for the narrowing of margins of time: (1) less expensive transportation options, (2) private importers moving to international connections for credit which are able to offer lower rates and cheaper financing, (3) a concentration in international fertilizer distributors enabling economies of scope and cost savings, and (4) increased competition at the local distribution level since the mid-1990s.

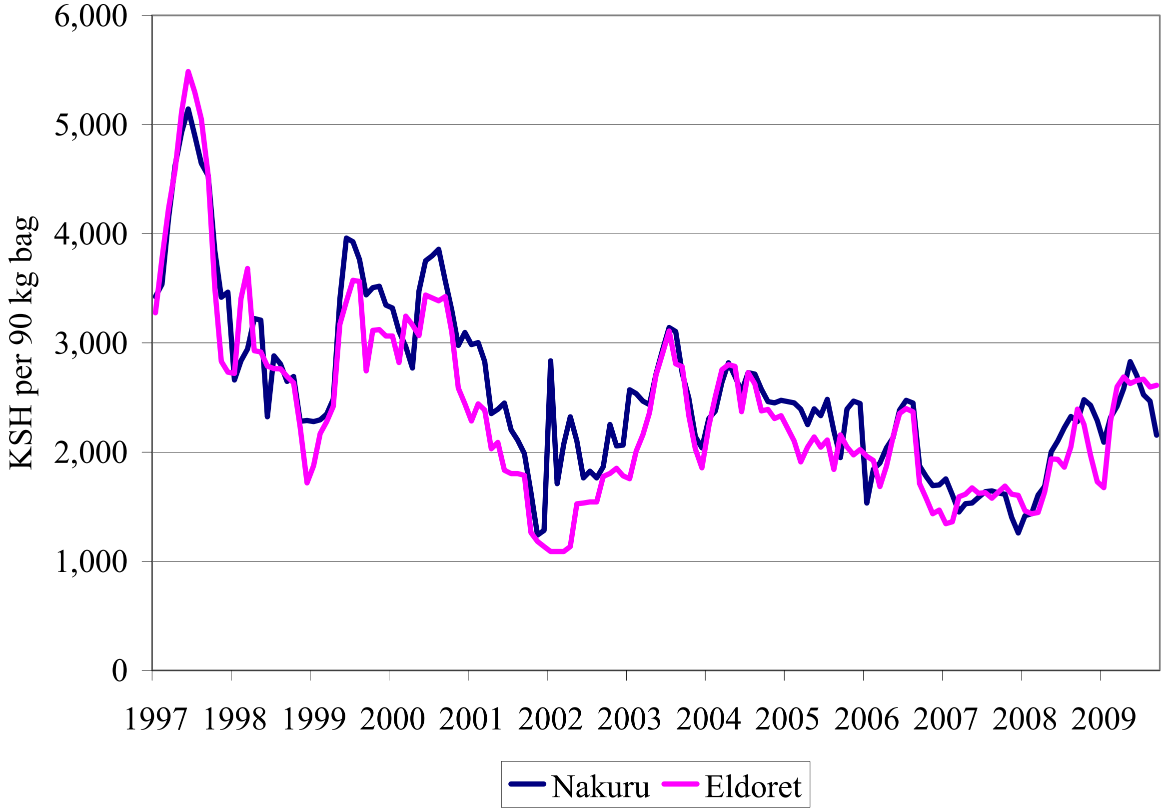
Figure 2: Price of DAP at Mombasa and Nakuru (in 2009 prices)



Source: Prices from Ministry of Agriculture in Kenya. Mombasa prices represent cif. Nakuru prices represent those at wholesale market. CPI from the Kenya National Bureau of Statistics.

Taken together, fertilizer consumption has increased while fertilizer prices have fallen, despite the price shock in 2007-2008. A hypothesis, then, would be that with a reduction in fertilizer prices over time, fertilizer application has become more profitable, leading to the observed increase in use. Fertilizer prices, however, are only one part of the economic profitability calculation; the price of output is just as important in assessing the incentive to use fertilizer. Figure 3 shows the real price of maize grain at two major wholesale markets in Kenya (Nakuru and Eldoret), measured monthly. This graph shows that, like fertilizer prices, maize prices also have fallen over time, even with considerable price spikes in 2000, 2004 and 2009.

Figure 3: Real maize grain prices in major maize producing areas (in 2009 prices)



Source: Maize prices from Market Information Bureau, Ministry of Agriculture Kenya. The CPI from the Kenya National Bureau of Statistics.

With a downward trend in both fertilizer and maize prices, this calls into question how relative prices and, therefore, incentives to use fertilizer have changed over time. A number of previously highlighted studies investigated fertilizer profitability using relative prices as part of their profitability measure but focused on a small geographic area over a relatively short period of time. This analysis will expand on their work by using variation across space and time to look at fertilizer profitability under a number of relative price scenarios to better understand how the national-level trends described here translate into community-level trends.

Chapter 2: Conceptual Framework and Methodology

In this chapter, I detail how I conceptually frame fertilizer use in the context of economic profitability. With a conceptual understanding, I then explain how I methodologically put these ideas into practice for completing the analysis in subsequent chapters. The sequence of this analysis is similar to work previously done by Hassan et al. (1998) in Kenya, combining research at agricultural experiment stations with household surveys. This analysis goes beyond that work by focusing specifically on household level survey data, using a data set covering a longer period of time and updated with more current information, and looking specifically at fertilizer use profitability instead of simply maize response and household use.

1. Conceptual Framework

Households in Kenya typically function as multiproduct firms, deriving income from the production of various crops and often a range of off-farm activities. I assume households are optimizers subject to constraints across all activities. With respect to agricultural production, which accounts for a large percentage of potential income of rural households in Kenya, I assume households optimize not only over all activities, but also at the field level. Maize production is generally one of the most important household activities in Kenya given maize is the overwhelming staple in the Kenyan diet and the crop most often found on farms across the country. Given the importance of maize in the Kenyan production system and the fact that available data is specific to maize, this analysis focuses on the maize enterprise or, more specifically, maize fields.

The yield Y on maize field i from household j during year t is a function of a vector of physical inputs x and characteristics of the household z:

$$Y_{ijt} = f(x_{kijt}, z_{kijt}, \mu_{ijt}) \quad (1)$$

where the vector x is comprised of both inputs chosen by the household (e.g., fertilizer, seed) and the agro-ecological conditions of the field in question (e.g., soil attributes, rainfall); the vector z includes those characteristics of the household that likely contribute to yield (e.g., skill level of the production manager); and μ is the error term comprised of unobservable characteristics of the production system that affect yield with or without knowledge of the household.

With an accurately specified production function, I can calculate the contribution of each input to maize yield and, subsequently, combine with relative input to output prices to calculate the profitability of input use via marginal and average value cost ratios (MVCRs and AVCRs).[1] While similar in their derivation, MVCRs and AVCRs allow us to understand different facets of input profitability, making it necessary to analyze both. Depending on the form of the production function and the magnitude of the estimated coefficients, these numbers could be very different or essentially the same. Marginal and average value cost ratios are calculated as follows:[2]

$$MVCR_{fijt} = \frac{p_{yt} * MPP_{fijt}}{w_{fijt}} \quad (2)$$

$$AVCR_{fijt} = \frac{p_{yt} * APP_{fijt}}{w_{fijt}} \quad (3)$$

where p_y is the output price of maize, w_f the input price of fertilizer, MPP_f the marginal physical product of fertilizer, and APP_f the average physical product of fertilizer. The marginal physical

[1] These measures assume that there are no other major additional costs to the farmer in using fertilizer besides the cost of the fertilizer itself.

[2] These equations require independence between the included terms, which is a reasonable first order approximation unless markets are very localized.

product (MPP) of an input is derived from the production function Y_{ijt} by taking its first derivative with respect to that input x_{ijt} and describes how much extra output can be produced by using one additional unit of a given input, all else held constant.:

$$MPP_{xijt} = \partial Y_{ijt}/\partial x_{ijt} \tag{4}$$

Generally, the average physical product (APP) is calculated as output divided by the amount of variable input used:

$$APP_{xijt} = Y_{ijt}/x_{ijt} \tag{5}$$

However, one can also conceptualize and calculate average product slightly differently:

$$APP_{xijt} = \frac{\hat{Y}^{w} - \hat{Y}^{wo}}{x_{ijt}} \tag{6}$$

where $\hat{Y}^{w}$ is the predicted yield with the input x_{ijt}, $\hat{Y}^{wo}$ is the predicted yield without x_{ijt}, and x_{ijt} is the amount of the input used. This method of calculating the average product describes the gain in yield per unit of an input relative to not using any of that input and is used in this analysis.

An AVCR of greater than one means that a risk neutral household could increase its income as a result of fertilizer use (i.e., the average gain per unit); an MVCR of greater than one indicates income would be increased with an increase in the rate of fertilizer application (i.e., the gain to the last unit). As such, the risk neutral household makes decisions regarding fertilizer application—both whether or not to use and, if so, how much—with the following two rules:

$$MVCR_{fijt} \geq 1 \tag{7}$$

$$AVCR_{fijt} \geq 1 \tag{8}$$

However, given the fact that households in Kenya may be risk averse, I include a risk premium ρ in the set up (e.g., Anderson et al. 1977). An MVCR of two (meaning a risk premium of one) has been used in the literature (e.g., Xu et al. 2009 in Zambia; Sauer and Tchale 2009 in Malawi; Bationo et al. 1992 in Niger) dating back to work by the FAO (1975) in order to better accommodate risk and uncertainty, adjust for the many unobserved costs associated with fertilizer use, and serve as an approximation for the rate at which fertilizer is profitable *enough* for farmers to want to use it, generally for the first time (see Kelly 2005). Furthermore, because farmers make the decision to use fertilizer before all relevant variables are known, I estimate *expected* marginal and average value cost ratios:

$$E(MVCR_{fijt}) = \frac{E(p_{yt}) * E(MPP_{xijt})}{w_{fijt}} \tag{9}$$

$$E(AVCR_{fijt}) = \frac{E(p_{yt}) * E(APP_{xijt})}{w_{fijt}} \tag{10}$$

and related decision rules:

$$E(MVCR_{fijt}) \geq 1 + \rho \tag{11}$$

$$E(AVCR_{fijt}) \geq 1 + \rho \tag{12}$$

The key objective of this thesis is to understand whether or not farmers are complying with the decision rules described in equations 11 and 12, meaning estimating whether or not farmers are using fertilizer on maize fields when it is economically profitable for them to do so within reasonable bounds of risk and uncertainty. When farmers make the choice contrary to these rules, understanding the most likely reasons for this choice is important for improving fertilizer application recommendations and development programs.

2. Production Function and Econometric Techniques

In Chapter 4, I estimate a maize yield response model (i.e., production function) as described in equation 1 which forms the basis of estimation and subsequent analysis on fertilizer profitability and use. As such, the functional form of choice is critically important to accurately describing the production environment in which Kenyan smallholder farmers operate and producing unbiased estimates of the parameters. In their review of over twenty functional forms, Griffin et al. (1987) detail a set of criteria for choosing one of the many established forms of production functions including (1) consideration of the maintained hypotheses, (2) constraints to estimation including data availability and properties, (3) goodness-of-fit and general data conformity and (4) the application of results.

Within the literature of yield response to fertilizer application, there are several camps of opinion regarding the most appropriate functional form given both theoretical considerations and observed complementarily with biological production processes. The quadratic (or higher order polynomial) functional form is often employed because it allows for concavity and diminishing returns. Aggregation across space tends to result in nonlinear responses that can be approximated by a quadratic at the field level even when there are linear to plateau relationships in some areas, making it a good first order approximation to many functional forms. This is particularly true when there is substantial heterogeneity across fields. There is a literature, however, that points to the many shortcomings of polynomials including the fact that they often overestimate yield and optimal fertilizer use recommendations and fail to consider minimum levels of inputs necessary for growth.

These criticisms led to development of the von Liebig functional form and subsequent analyses by Ackello-Ogutu et al. (1985) and Grimm et al. (1987) showing that von Liebig forms consistently produced superior estimates to the polynomial. Von Liebig models assume that yield will be constrained by the most limiting input but assume a lack of complementarity between input types, meaning a fixed proportion is required for plant growth. Another adaptation of the von Liebig model is the linear response and plateau (or LRP) which, like the von Liebig, considers restrictions by limiting inputs and does not allow for substitution and, unlike the von Liebig, forces an upper bound to yield through the use of a plateau.

While the idea of a limiting input is useful, the empirics only allow one input to be specified as the limiting one. This makes sense in an experimental context where inputs and conditions are closely controlled, but perhaps not when considering a heterogeneous mix of farmers and growing conditions (i.e., rainfall could be the limiting input in one area while available nitrogen in another). When considering heterogeneous conditions, Berck and Helfand (1990) show that the polynomial and LRP approximations essentially converge, making the quadratic a viable alternative to the von Liebig and LRP models. Similarly, when testing the goodness-of-fit of von Liebig models, Berck et al. (2000) find that these models generally do not fit the data well and that actual estimation does not yield the right angel isoquants described in its derivation.

Relying on these findings and other studies looking at smallholder production systems with similar attributes (e.g., Traxler and Byerlee 1993; Kouka et al. 1995), I estimate a modified quadratic production function of the following form:

$$Y_{ijt} = \Sigma\,(x_{kijt} + x_{kijt}^{2}) + \Sigma\,(z_{kijt} + z_{kijt}^{2}) + \Sigma\, x_{kijt}\, z_{kijt} \qquad (13)$$

where field *i* of household *j* during year *t* is a function of each input *k* from the vectors x and z. Each input has a linear and squared term and is interacted with other inputs. The quadratic form I use is “modified” because not all possible interactions are estimated; instead, only those with conceptual significance are included in order to maximize available degrees of freedom. A level (and linear) production function, as estimated here, is better able to deal with zeros in the dependent and independent variables, as I would expect there to be in a heterogeneous production environment such as Kenya. This analysis builds on a model originally constructed by Ariga (forthcoming), which uses a subset of the same household level panel data from smallholder farmers across Kenya.

2.1. Estimation techniques

Even with a rich household dataset with a large number of observable physical and environmental inputs, there is good reason to believe that some important variables in determining yield are unobserved (e.g., skill level of the farm manager). If the unobserved variable c is uncorrelated with any of the other inputs from the vectors x or z, then consistent estimators can be recovered using pooled ordinary least squares (OLS) estimation. If c is correlated with any k in x or z, then a different estimation technique is necessary (Wooldridge 2010). In this production function, there is reason to believe that managerial skill, for example, is correlated with both the amount of fertilizer applied and the amount of seed applied, likely among others. Another technique is, therefore, necessary for recovering unbiased and consistent estimates of the parameters.

Using a panel dataset enables consideration of a variety of techniques to control for unobserved heterogeneity, including random effects (RE), fixed effects (FE), and correlated random effects (CRE). While commonly used in panel data analysis, the main limitation of the random effects (RE) estimator is that it relies on the assumption that unobserved heterogeneity is uncorrelated with any of the observed independent variables. This assumption is likely too strong for this context. The fixed effects (FE) method relaxes this assumption, but does not allow estimation of coefficients on time invariant parameters, some of which are of interest in this thesis (e.g., soil type). More appropriate for this analysis, then, is the correlated random effects (CRE) estimator which both allows for correlation between the unobserved omitted variable c and included explanatory variables *k* in x and z and enables estimation of the effects of time invariant variables. CRE models use a device modeled by Mundlak (1978) and Chamberlain (1980) which, instead of treating the omitted variable as a parameter to estimate, allows modeling the distribution of the omitted variable conditional on the means of the strictly exogenous variables:

$$c_j = \tau + \overline{x}_k \gamma + a_{ijt} \quad (14)$$

where $\overline{x}_k$ is a vector of average values of each input x_k at the household level *j* across all waves of the panel. The production function with added Mundlak-Chamberlain device is equivalent to the household fixed effects estimator in this context because the model is linear.

I estimate the CRE model using pooled ordinary least squares (OLS). Using pooled OLS makes estimates more robust but potentially less efficient than maximum likelihood estimation (MLE). MLE, however, requires an extra assumption about the structure of the variance matrix which, if not true, produces biased estimates of the parameters.

2.2. Identification assumptions and strategies

Several assumptions are made to help with identification of the parameters. First of all, all fields are estimated together in the same model. I remove some of the heterogeneity across fields and households through (1) the use of the Mundlak-Chamberlain device, which controls for unobserved heterogeneity at the household level, and (2) the use of conditioning variables, which allows me to control for the environmental context that contributes to differences in yield response across households. After removing these important sources of heterogeneity, I assume the remainder of the variation in input and output levels, and that which is used to estimate the parameters of interest, comes from differences across space and time in (1) economic incentives (i.e., relative input and output prices), (2) constraints in the system (i.e., fertilizer availability), and (3) household preferences.

Secondly, I assume that households make the decision to use fertilizer at the beginning of the season before exogenous shocks occur (e.g., pest or parasitic striga infestation). Similarly, households have expectations about input responsiveness and yields based on previous experience, meaning their production function is known, to a large extent. So, while inputs are not randomly allocated (as they could be, for example, in an experimental context), I assume that households make both cropping and input decisions with a good sense of the production system unique to them and the field in question.

3. Computing Fertilizer Profitability

In Chapter 5, I compute the profitability of fertilizer use as described in the conceptual framework. Per equations 9 and 10, there are three important values that comprise the expected MVCR and AVCR calculations: (1) the marginal physical product (MPP) or average physical product (APP) of fertilizer, (2) the output price of maize, and (3) the input price of fertilizer. The expected MPP and APP of fertilizer are calculated using coefficient estimates from an accurately specified production function, disaggregated to the district and soil group level, as described further in Chapter 4. The output prices of maize and input prices of fertilizer are calculated as district level averages of all values observed in the data set, described further in Chapter 5.

The price of fertilizer is not necessarily limited to its market price. There are several fixed costs associated with buying fertilizer which involve both transaction costs, the costs associated with partaking in an economic exchange (Coase 1960; Williamson 1979), and transportation costs, the costs associated with moving the fertilizer from its purchase location to the farm. In the case of fertilizer acquisition, transaction costs can include the search cost for identifying price of the input, information costs associated with knowing what amount of an input to apply, and the opportunity cost of work time foregone in transport. The significance of transactions and transport costs in limiting farmers ability to participate in markets—both input and output—is well-established in the literature (de Janvry et al. 1991; Key et al. 2000; Bellemare and Barrett 2006) and has been used to explain why input adoption may be lower than expected (Morris et al. 2007; Winter-Nelson and Temu 2005 in Tanzania). Most related, Alene et al. (2008) focus on transactions costs in their assessment of fertilizer use in Kenya, finding that high transactions costs can have significant negative effects on market participation but that institutional innovations can often overcome them. In perhaps the first estimation of the size of transactions costs, Renkow et al. (2004) use data from maize farmers in Kenya to find an average ad valorem tax equivalent of transactions cost to be about 15 percent. While transport costs are relatively straightforward to estimate, a full set of transaction costs is rarely observable, making the true fixed cost of acquisition impossible to discern. Instead, I compute and use a transport cost of acquiring fertilizer as an approximation of the full set of fixed costs.

Because a large portion of households in this data set are net buyers of maize (as opposed to net sellers), I also take and use district-averaged maize buying prices as an additional measure of the opportunity cost of producing maize. The fact that a majority of households, even in agriculturally dominant areas, are net buyers has been well-documented by other researchers with respect to all of SSA (e.g., Christiaensen and Demery 2007) and Kenya specifically (e.g., Jayne et al. 2001). Furthermore, a relatively small number of farming households comprise the total marketable surplus of maize in the country. Jayne et al. (2001) found that 10 percent of small scale farmers produced 74 percent of the maize sold by the small scale maize sector. Because the opportunity cost of maize production is different between net buyers and sellers, I estimate fertilizer profitability with both prices for comparison.

Using the array of values described above, I consider five scenarios of relative prices (i.e., selling and buying price of maize paired with the market and market plus transport cost of fertilizer) to look at how real changes in the relative price ratio might change when and where fertilizer use is profitable. To my knowledge, this is the first attempt not only to look at various profitability scenarios that mimic the actual and varied market conditions of farmers, but also to seriously consider the effects of observed transport costs in the profitability computation.

Then, because MVCRs and AVCRs are measures of relatively profitability (i.e., use the relative prices of input to output), I compute an additional measure of absolutely profitability, measuring the net gain in revenue to the last unit of fertilizer used. With a reduction in both fertilizer and maize prices over time (see Figure 2 and Figure 3), absolute profitability provides a better sense of how changes in absolute prices affect profitability, even when relative prices are fairly constant. This value is computed as:

$$\text{net gain to last unit of fertilizer} = E(MPP_{xijt}) * E(p_{yt}) - w_{fijt} \qquad (15)$$

I compare this absolute measure to the relative measures discussed in earlier in the chapter for comparison.

4. Comparing Fertilizer Profitability with Observed Use Decisions

In Chapter 6, I seek to establish if and where a gap exists between where fertilizer is profitable and where farmers are using it and, furthermore, if there is room for profitable expansion of fertilizer application rates. First, I compare estimated fertilizer profitability measures from the previous chapter with observed use patterns. Using computed profitability levels by year from the previous chapter, I use descriptive statistics to investigate if there is a gap between where it is profitable to use fertilizer and where households are already found to be using it.

Then, I return to the production function estimates and compute the economically optimal amount of fertilizer application at the district level by soil type and compare calculated optimal levels to observed use levels. Under profit maximization conditions and risk neutral behavior, the economically optimal amount of input use is where the marginal value product (MVP) is equal to the marginal factor cost (MFC), which is equivalent to where the marginal value cost ratio (MVCR) equals one (see equation 9). However, because households may be risk averse in their decision to use fertilizer, I also solve for the rate of nitrogen where the marginal value cost ratio (MVCR) equals two for comparison.

Using both of these computed values, I measure the size of the "gap" between optimal and observed fertilizer application rates. This "gap" represents the constraints that limit farmer fertilizer use, technical inefficiency and the interactions between the two (Kumbhakar and Baushan 2009; Komicha and Ohlmer 2007), meaning its interpretation should be carefully considered. The size of the estimated gap will provide further evidence to or against the claim that farmers in Kenya currently are using fertilizer far below optimal and profitable levels.

Then, to return to the theme of absolute profitability, I compute the revenue added from application of fertilizer, both at current use levels and estimated optimal ones for comparison. This measure will provide further evidence for or against the claim that households are under-utilizing fertilizer where profitable. I compute the overall gain in revenue from fertilizer application as follows:

$$\text{net gain to total fertilizer application} = [E(Y^F) - E(Y^{NF})] * E(p_{yt}) - x_{ijt} * w_{fijt} \qquad (16)$$

where Y^F is yield with fertilizer application (at whichever level defined) and Y^{NF} is yield without fertilizer application. This measure provides insight into how current levels of fertilizer use and calculated optimal ones contribute to income levels of smallholder farmers. The "gap" between these values provides further insight into the absolute gains to fertilizer yet to be had by farmers.

5. Understanding the Decision to Use Fertilizer Where Profitable

In the presence of a gap between where fertilizer use is profitable and where farmers are using it (in any amount), I seek to understand the reasons this might be the case in Chapter 7. After limiting my sample to those areas where fertilizer use is estimated to be profitable, I investigate the constraints to using fertilizer both qualitatively, using stated reasons from households about why they choose not to use fertilizer on maize in a given year, and quantitatively, using regression analysis.

The most appropriate model specification includes a binary dependent variable, which takes a value of one when fertilizer is used and zero otherwise. There are a range of models that can be employed to study binary dependent variables, among them the linear probability model (LPM), probit model and logit model. Wooldridge (2009a) explains the shortcomings of the linear probability model, which uses standard ordinary least squares (OLS) as an estimator. The largest drawback is the lack of restriction on the dependent variable, which means predicted values can be less than zero or greater than one, neither of which are options in the set of actual possibilities. Two non-linear models were created to confine predicted values between zero and one: probit and logit models. Binary response models of these types take the following form:

$$P(y=1|x)=G(\beta_o+\beta_i x_i) \tag{17}$$

where P is the probability of dichotomous outcome y, which is dependent on a full set of explanatory variables x_i. In order for the predicted probability to fall between zero and one, the function G must take a particular form: the probit model assumes a standard normal distribution while the logit model assumes a logistic distribution. Otherwise, the two models are identical. Given the non-linearity of probit and logit models, the estimated coefficients cannot be interpreted like a linear model. Instead, partial effects, which transform the coefficients into their linear equivalents, can be calculated for ease of interpretation and comparing with the LPM equivalent. Given the non-linearity, probit and logit models are estimated using maximum likelihood estimation (MLE). Moreover, unlike the production function, I do not control for unobserved household heterogeneity in this model. While the Mundlak-Chamberlain device can be used in non-linear regression, there is no major benefit in estimating within-household variation (as would be the case with a CRE or FE model) given that several of the variables used in the regression have very little variation at the household level. Instead, I estimate a pooled model at the field level and control for a range of field, household, village and district level variables that likely influence the fertilizer use decision.

Adoption of improved technology (of which fertilizer is included) is a well-studied topic in the international development literature. Binary dependent variable models, generally taking the form of the more sophisticated probit and logit models, are commonly employed by researchers looking at fertilizer use decisions ("adoption") across Africa (e.g., Nkamleu 2000 on Cameroon; Daramola 1989 on Nigeria; Kebede et al. 1990 on Ethiopia) and specifically in Kenya (e.g., Waithaka et al. 2007; Ouma et al. 2006; Olwande et al. 2009). Throughout the literature, both theoretical and empirical, academics and practitioners alike propose a litany of reasons for potential non-adoption of a given technology (see Feder et al. 1985 for a survey). What makes this analysis different from most technology adoption studies, though, is the population of interest. Most studies look at a sample from an entire population to understand the differences between adopters and non-adopters; I, however, confine my population to those households where fertilizer use is estimated to be profitable using profitability calculations from the earlier part of this thesis. In doing so, I am able to exclude from the analysis households for which fertilizer use is unprofitable and therefore focus on the factors associated with constraints within environments where fertilizer use contributes to increases in household income.

Chapter 3: Data, Sample Selection and Summary Statistics

In this chapter, I describe the data used in this analysis and the sample I draw from it. Then, as a prelude to econometric analysis, I look at summary statistics on household level changes in fertilizer use over time using households in the described data set. A descriptive understanding of changes over time will guide interpretation and utility of econometric findings.

1. Data

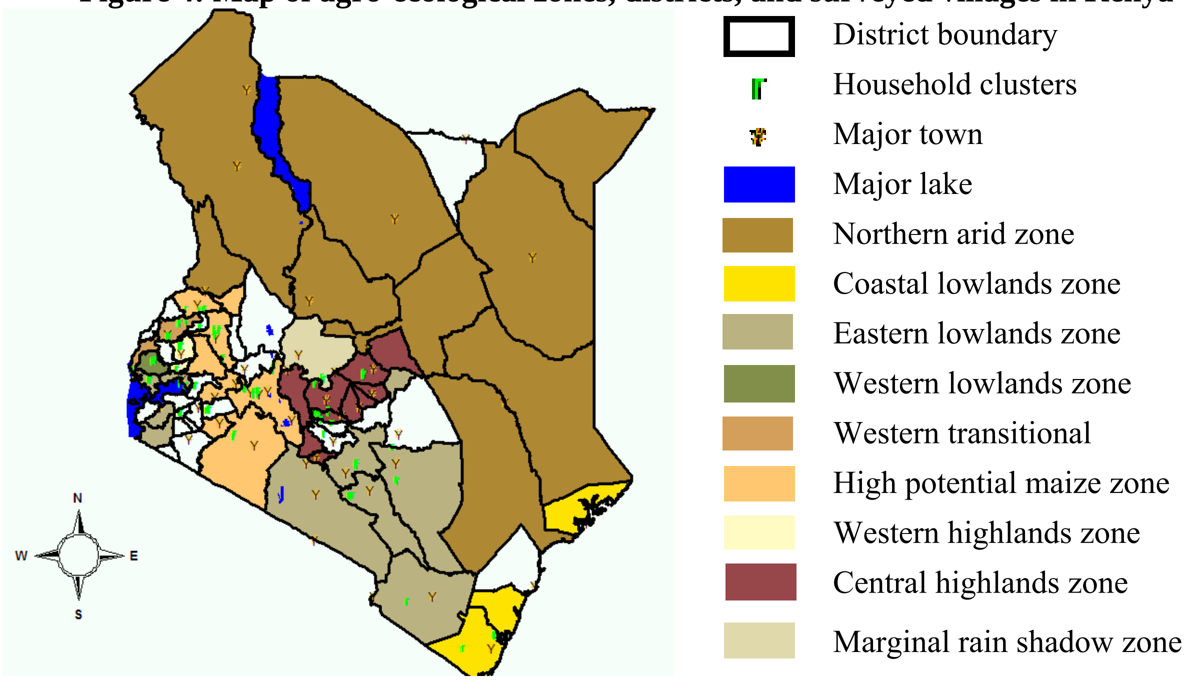
The data used in the analysis comes from Egerton University's nationwide Tegemeo Rural Household Survey for the years 1997, 2000, 2004, 2007 and 2010. Households are asked a range of questions about their agricultural activities, other sources of income, and demographics. The surveys geographically cover 24 administrative districts, 39 divisions and 120 villages where standard proportional sampling using census data for rural divisions of the country formed the basis of extraction of the sample households. The panel started with 1500 households but, due to attrition, 1243 are present through the final 2010 panel.[3] Supplemental data on yearly rainfall levels comes from the National Weather Service Climate Prediction Center (CPC) as a part of their Famine Early Warning System (FEWS) project; soil data comes from the Kenya Soil Survey and the Ministry of Agriculture from data originally collected in 1980.

Throughout this text, I most often refer to the agro-ecological zone (AEZ) in which a particular household or village is located. These agro-ecological zones are defined based on similarities in agricultural, ecological, and environmental conditions. Given interest in estimating maize yield response, which is a biological process influenced by these factors, these zones are

[3] 1500 are present in 1997, 1407 in 2000 (6.2 percent attrition rate), 1324 in 2004 (5.9 percent attrition rate), 1275 in 2007 (3.7 percent attrition rate), and 1243 in 2010 (2.5 percent attrition rate) for an overall attrition rate of 17.1 percent.

my preferred geographic unit of grouping households. Figure 4 graphically shows where these zones are located and how districts (i.e., administrative units) correspond.

Figure 4: Map of agro-ecological zones, districts, and surveyed villages in Kenya



Note: The Northern Arid zone, where agriculture is not as important, is the only zone not used in this analysis.

2. Sample Selection

From this data set, I narrow my focus to maize fields, the unit of analysis. Unlike many studies that average across fields to the household level, I keep the maize field my unit of analysis throughout. I limit the sample to fields that meet the following criteria: (1) have maize and no more than six other crops,[4] (2) maize is not produced alongside a major cash crop (i.e., tea, sisal, rice, pyrethrum, cotton), and (3) maize constitutes at least 25 percent of the calculated

[4] Six crops may seem like a lot, however the detail achieved in data collection indicates otherwise. For example, many Kenyan households choose to grow maize and beans together, often with a small amount of sukuma wiki (kale). Neither the beans nor the sukuma wiki contribute to lower maize yields because maize still constitutes a large portion of the field. Avocado and banana trees may line the perimeter of the field while pumpkin may be used as a cover crop. So, while defined as "intercropped," even fields with several other crops, maize still is the overwhelming dominant crop on the field. See Appendix 1 for more examples.

potential revenues from the field. Given less than 10 percent of fields are maize monocropped, this criterion allows a larger number of fields to be considered while still focusing on maize as the crop of interest. On average across years, about 75 percent of households have one maize field per year, 20 percent have two, and the remaining 5 percent have three or more. Table 1 shows the percentage of all fields in this data set that are classified as either maize monocropped or maize intercropped in each survey year per the above requirements.

Table 1: Percentage of all fields categorized as maize (non-maize fields excluded)

		1997	2000	2004	2007	2010
Coastal Lowlands	Maize mono	14.6	5.8	6.5	10.0	5.1
	Maize inter	35.4	25.8	28.2	26.3	45.5
Eastern Lowlands	Maize mono	7.4	8.9	7.4	8.4	6.6
	Maize inter	38.6	25.2	30.8	33.2	32.8
Western Lowlands	Maize mono	5.8	5.9	3.2	1.6	4.2
	Maize inter	40.5	42.2	27.9	27.1	32.7
Western Transitional	Maize mono	5.2	1.9	2.6	3.6	2.1
	Maize inter	20.5	17.9	17.3	18.8	20.5
High Potential Maize	Maize mono	7.6	2.4	3.4	6.2	4.9
	Maize inter	25.2	15.5	14.8	17.2	19.5
Western Highlands	Maize mono	7.6	2.4	1.3	0.7	1.5
	Maize inter	30.6	31.5	26.4	24.7	27.1
Central Highlands	Maize mono	4.2	3.2	3.3	3.3	2.3
	Maize inter	22.6	14.3	16.1	16.6	20.4
Marginal Rain Shadow	Maize mono	0.0	4.2	0.4	1.2	3.3
	Maize inter	53.2	38.2	31.8	42.7	36.6
Total sample	Maize mono	6.2	4.0	3.4	4.2	3.6
	Maize inter	28.5	21.5	20.7	21.7	24.9

Note: Includes fields from all 1243 households in balanced panel. Maize monocropped fields are defined as having only maize. Maize intercropped fields include those with maize and at most six other crops. See text for additional criterion.

Given I do not observe the same field over time and that the composition and number of fields at the household level can vary between survey years, the resulting panel is unbalanced. Wooldridge (2009) shows that correlated random effects (CRE) can be employed with unbalanced panels in linear models, such as the quadratic production function estimated here, to produce similarly unbiased results. Therefore, instead of sampling from the 1243 households in the five year balanced panel, I include maize fields from any household interviewed in each of the five waves. Moreover, while some households might have multiple maize fields, not all households produce maize or have qualifying maize fields. Under the previously stated assumption that households choose their cropping pattern with optimization in mind, the fact that some households from the panel do not appear in estimation should not bias my estimates, particularly given my focus on the maize enterprise. Because I choose a specific population of fields from a random sample, the resulting data set is representative of Kenyan maize producing regions. Additionally, because Mundlak-Chamberlain is used to control for unobserved heterogeneity at the household level in the production function and variation in the explanatory variables is necessary for the household level averages to be a viable control, a household must have maize fields in at least three of the five survey years to be considered in the model. The reason for choosing three years is that is allows a sufficient number of observations for creating an average not specific to or skewed by one year without limiting the sample size too much. Households that do not meet this criterion are dropped from production function estimation.

Table 2 shows the number of households and observations (one for each field and year) by zone and district used to estimate the production function, bringing the total number of households to 906 (4714 fields). Notice that observations in the Coastal Lowlands and Marginal Rain Shadow are dropped from analysis; this is due to the fact that agricultural conditions in these zones are very different from the others, making it difficult to get good estimates. This and other production function specifics are described in greater detail in Chapter 4. The final column in Table 2 shows the number of households that remain for the binary fertilizer use decision analysis. Given interest in understanding the fertilizer use decision were I estimate fertilizer use to be profitable, I limit my sample to only those areas where fertilizer use is, on average, profitable. Furthermore, given both data limitations and the desire to focus on fertilizer use in the recent past, only observations from the last four survey years (1997 excluded) are used in this sample. Like the production function, the unit of analysis remains the maize field. All of these adjustments to the production function sample bring the binary fertilizer use sample size to 882 households (3521 fields). More specific sample selection issues for the binary fertilizer use model can be found in Chapter 7.

Table 2: Distribution of households (and fields) used in analysis

Agro-ecological zones	Districts	Original panel	Balanced panel	Production function sample	Fertilizer use model sample
Coastal Lowlands	Kilifi, Kwale	80	74	0	0
Eastern Lowlands	Machakos, Mwingi, Makueni, Kitui, Taita-Taveta	166	141	103 (528)	103 (447)
Western Lowlands	Kisumu, Siaya	188	149	41 (248)	41 (206)
Western Transitional	Bungoma (lower elevation), Kakamega (lower elevation)	172	145	154 (822)	154 (670)
High Potential Maize Zone	Kakamega (upper elevation), Bungoma (upper elevation), Trans Nzoia, Uasin Gishu, Bomet, Nakuru, Narok	411	331	341 (1841)	332 (1262)
Western Highlands	Vihiga, Kisii	156	128	135 (738)	120 (517)
Central Highlands	Nyeri, Muranga, Meru	268	241	132 (537)	132 (419)
Marginal Rain Shadow	Laikipia	59	34	0	0
Total sample		1500	1243	906 (4714)	882 (3521)

Note: See text in this chapter for discussion on criteria used to determine which households and maize fields are included in estimation. See Chapter 4 for further discussion on the production function sample. See Chapter 7 for more information on the binary fertilizer use model sample.

3. Summary Statistics on Household Level Fertilizer Trends

While the national level trends described in Chapter 1 are helpful in providing context, conditions across Kenya are tremendously varied and the story described there may not necessarily hold for all locations. Before moving to econometric analysis, this section provides insight into how fertilizer use has changed over time in Kenya using households from this dataset.

Table 3: Percent of fields where fertilizer was applied in any amount by type of field

		1997	2000	2004	2007	2010
Coastal Lowlands	Any field	2	3	4	7	8
	Maize field	3	5	5	11	17
Eastern Lowlands	Any field	21	18	24	33	27
	Maize field	23	24	41	40	51
Western Lowlands	Any field	3	4	4	9	10
	Maize field	2	3	5	12	13
Western Transitional	Any field	20	29	31	39	36
	Maize field	38	63	74	80	77
High Potential Maize	Any field	53	43	48	51	47
	Maize field	78	87	87	90	89
Western Highlands	Any field	45	52	47	45	44
	Maize field	72	88	91	93	94
Central Highlands	Any field	57	59	51	57	63
	Maize field	87	86	86	90	84
Marginal Rain Shadow	Any field	14	15	11	23	11
	Maize field	4	4	4	13	6
Total sample	Any field	38	37	37	41	40
	Maize field	52	56	64	68	67

Note: Fields are identified by the households during data collection. All available fields, not just those from the balanced panel, are used here.

Table 3 shows how the percent of households using fertilizer in each zone has changed over the survey years. The first row in each agro-ecological zone shows the percentage of households that used fertilizer on any crop or field while the second row is specific to application on maize fields. Notice how percentages and changes vary considerably across zones. In the higher potential maize regions (i.e., Western Transitional, High Potential Maize Zone, Western Highlands and Central Highlands), over 70 percent of households currently use fertilizer with some zones closer to 95 percent. In the generally lower potential maize production areas (i.e., Coastal Lowlands, Eastern Lowlands, Western Lowlands and Marginal Rain Shadow), percentages are often much lower, although more varied. A higher portion of households in the Eastern Lowlands use fertilizer on maize (almost 50 percent currently) compared to the other areas where 10 to 20 percent is more common. These areas, however, have seen a doubling or more of households using fertilizer between 1997 and 2010.

Table 4: Mean kilograms per hectare of fertilizer applied to maize fields (excludes zeros)

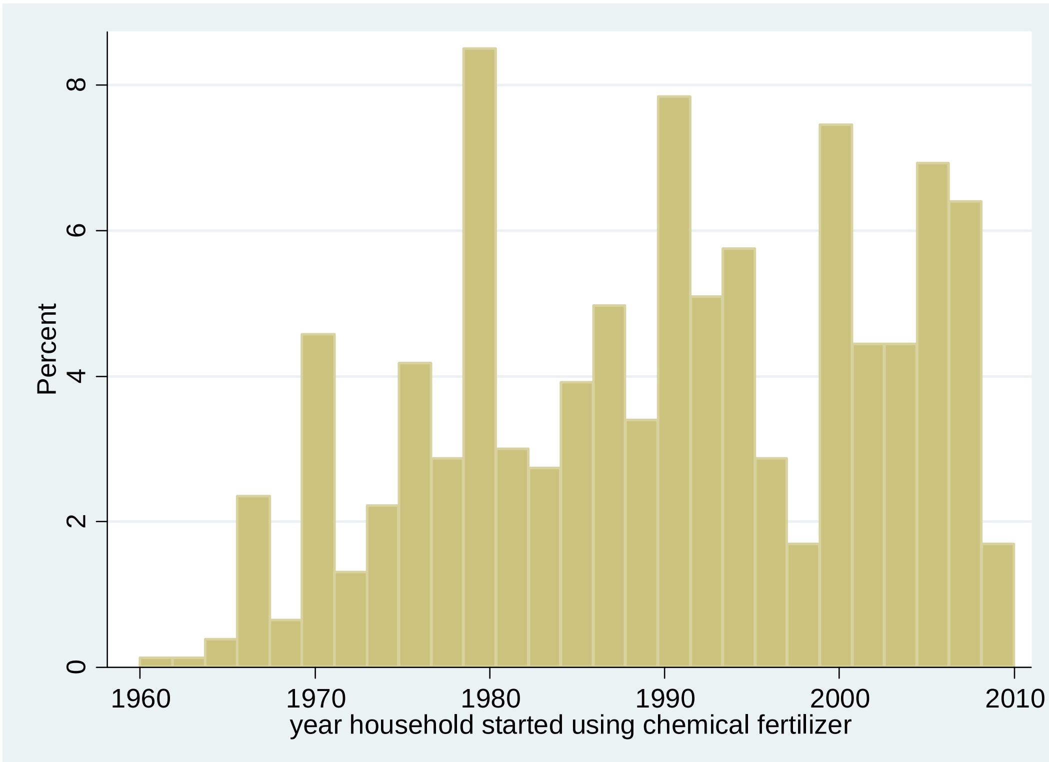
		1997	2000	2004	2007	2010
Coastal Lowlands	Monocrop	-	-	4.3	39.5	14.8
	Intercrop	2.5	24.3	6.2	42.2	44.2
Eastern Lowlands	Monocrop	40.3	32.2	61.7	42.5	139.9
	Intercrop	26.3	53.3	42.0	72.3	106.4
Western Lowlands	Monocrop	67.1	-	-	-	123.5
	Intercrop	85.6	43.1	46.2	52.1	122.9
Western Transitional	Monocrop	168.5	160.2	156.9	144.2	185.2
	Intercrop	137.8	120.8	154.9	177.9	180.8
High Potential Maize	Monocrop	168.5	167.3	210.8	211.3	209.3
	Intercrop	145.3	169.2	167.2	178.9	190.2
Western Highlands	Monocrop	78.3	89.5	30.2	150.9	129.5
	Intercrop	88.0	83.6	127.1	124.7	181.4
Central Highlands	Monocrop	131.6	131.6	132.3	120.1	137.7
	Intercrop	145.1	119.0	123.4	116.8	151.0
Marginal Rain Shadow	Monocrop	-	-	-	-	-
	Intercrop	-	-	-	63.9	123.5
Total sample	Monocrop	136.1	143.4	157.2	179.0	185.1
	Intercrop	130.9	134.1	139.6	148.6	169.5

Note: Includes all maize field observations. Extreme values of fertilizer application eliminated. Refer to Table 7 for how these values translate into nitrogen and phosphorous application rates.

Another useful statistic is how much fertilizer households are applying and how that number changes over time. Table 4 shows the average kilograms per hectare of fertilizer applied to both monocropped and intercropped maize fields across all five waves of the panel. Again, these numbers show great diversity across Kenya. In general, monocropped maize fields receive less fertilizer than intercropped fields, meaning farmers choose to fertilize non-maize crops more heavily or often. In the high potential areas, monocropped fields are fertilized at a rate between

125-225 kilograms per hectare. Farmers in the Western Highlands fertilize at rates similar to those in the Western Lowlands, the former considered high potential and the latter low potential. Otherwise, in the Coastal Lowlands and Marginal Rain Shadow, fertilizer has been applied only in the recent past while the Eastern Lowlands has seen a tremendous increase in application rates since 1997.

Figure 5: Distribution of year in which household started using inorganic fertilizer



Note: Only includes households surveyed in 2010.

In 2010, all households were asked in what year they started using inorganic fertilizer on any crop, not just maize. The distribution of responses is found in Figure 5. Again, differences across agro-ecological zones are immense. On average, households in the High Potential Maize Zone and Central Highlands claim to have been using fertilizer for about 25 years as compared to about 10 in the Eastern and Western Lowlands. This history of diffusion closely follows how fertilizer came to exist in Kenya; it was first used only by European colonists on cash crops, who preferred growing conditions in the Rift Valley and surrounding highlands, then was taken up by Kenyan farmers on their own cash crops following independence in 1963 (Hassan et al. 1998). A government fertilizer subsidy coupled with the release of hybrid seeds further encouraged Kenyan farmers to start using fertilizer on their maize in the 1960s (Kimuyu et al. 1991). The lowlands areas, furthest from where fertilizer was initially introduced, were the last to start using fertilizer.

What these numbers mask, though, is the differences in conditions necessary for maize growth—namely soil type and fertility and rainfall amounts—and the profitability of using fertilizer given those conditions and input and output prices. While fertilizer application rates are much lower in some zones than in others, the maize yield response associated with fertilizer use in those areas might be such that using more fertilizer is not profitable or not profitable at the same levels of application. In the next chapter, I model maize production as a function of various inputs, fertilizer among them, to better understand the differences in fertilizer application rates, maize response and fertilizer profitability across Kenyan households.

Chapter 4: Maize Yield Response to Fertilizer Application

In this chapter, I estimate a maize yield response model to understand the contribution of inorganic fertilizer application (among other inputs) to maize yields. Section 1 describes the variables included in the yield response model; section 2 describes how households were grouped for estimation; section 3 describes the results of model testing; and section 4 describes the regression results and marginal and average products of fertilizer.

1. Description of Variables in the Yield Response Model

In this section, I discuss the variables used in the production function. Table 5 includes a complete list of those included and what they measure. The distribution of these variables over the entire sample and associate standard deviations can be found in the Appendix 5. Most inputs in the production process were collected at the field level; however, some are observed at the household, village, district, or zone level. The level of aggregation for each variable is described below. A number of missing and extreme values are dropped from the dataset prior to regression in order to limit the leverage of potentially erroneous observations. Field level observations are dropped if they satisfy any of the following conditions: (1) any missing value in the regressed variables, (2) plot size less than 0.06 hectares or greater than 7 hectares (due to the high likelihood of measurement error in input and output rates), (3) yield per hectare of greater than 9,700 kilograms, (4) maize seed per hectare of zero or greater than 60 kilograms, (5) nitrogen per hectare of greater than 120 kilograms or (6) phosphorous per hectare greater than 50 kilograms. These ranges were determined based on an understanding of reasonable values in the Kenyan context and government input recommendations.

Table 5: Description of variables included in production function

y		Output (yield)	Maize yield computed using Liu and Myers index
x	continuous	Nitrogen (N)	Nitrogen content of applied fertilizers (kg/hectare)
		Phosphorous (P)	Phosphorous content of applied fertilizers (kg/hectare)
		Seed (seed)	Seed rate (kg/hectare)
		Hectares (hect)	Number of hectares in given maize field
		Rainfall – moisture stress (rain)	Proportion of 20-day periods when rainfall was less than 40 mm during the main growing season
		Asset wealth (asset)	Value of assets at household level per hectare (proxy for household soil fertility and capital availability)
	dummy	Hybrid seed (hybrid)	1=new hybrid, 0=other seed (retained hybrid, OPV, local variety)
		Manure or compost (manure)	1=manure or compost applied to field, 0=none
		Legume intercrop (legume)	1=legume intercropped with maize; 0=none
	categorical	Crops per field (crop)	Number of crops included on field (range 1-7)
		FAO soil classification (FAO)	Type of soil as defined by the FAO soil classification system: Cambisols, Ferralsols, Phaeozems, Luvisols, Greyzems, Podzols, Regosols, Rankers
		Soil groups (soil)	Soils grouped into four based on above classification system: 1=volcanic, 2=high humus or highly productive, 3=Rankers with high sand, 4=Rankers with less sand
		Agro-ecological zone groups (zone)	Six agro-ecological zones grouped into three: 1=lowlands, 2=transitional and high potential, 3=highlands
		Years (year)	Each survey year included as a dummy
		Districts (dist)	Each district included as a dummy

Note: Terms in parenthesis in third column represent the variable names used in the text for simplicity. For more on the distribution of these variables, see Appendices 6 and 7.

1.1. Maize output: "revenue" yield index

Recall that a large portion of maize fields in this data set are intercropped, not monocropped (see Table 1). Therefore, in order to transform observed kilograms harvested of other crops into their maize output equivalents, an output index used by Liu and Myers (2009) of the following form is employed:

$$Y_{ijt} = \frac{\sum_n Y_{is} P_s}{P_m} \qquad (18)$$

where Y_{ijt} is the output index of field *i* at household *j* during year *t*, Y_{is} is the total kilograms harvested of crop *s* on field *i*, P_s is the market price of crop *s*, and P_m is the market price of maize. Note that for monocropped fields, the output index is simply total kilograms of maize harvested.[5] However, on intercropped fields, the output index resembles revenue and is conditional on the relative output prices and volume harvested of other crops. While the index creates a measure of field level revenue, I will refer to output throughout the text as "yield," for simplicity. Prices used in this computation are district level averages.[6] Even if the household did not sell its harvest of any particular crop, it will be valued at the level of those who did.[7] Examples of how this output index works for different field compositions (e.g., harvest amounts and relative price scenarios) are found in Appendix 1.

Table 6 shows computed maize yield per hectare. "Maize output" is generally higher on intercropped fields than monocropped ones, meaning the other crops planted on intercropped fields are either of higher value or have higher yielding capacity than maize. Variation in yield across time and geography is immense. Across the total sample, yield levels do not appear hugely different from one year to the next; however, at the zone level, differences are more

[5] Where both green and dry maize are harvested on the same field, the field is classified as monocropped. However, green maize and dry maize have different market values and therefore are considered separately in the yield computations. See Appendix 1 for an example.

[6] Where there are less than five data points per district, I use an average at the zone instead. If there are less than five data points at the zone level, then I use the national level price instead. This method seeks to cut down on over-weighting price observations where there is little market exchange of a particular crop in a given district.

[7] For some crops, this could mean over-valuing the crop relative to how the household views it. For example, if most households grow pumpkin for consumption and very few buy and sell, a household will see pumpkin as having very little monetary value but the district average might be relatively higher, reflecting the low supply of pumpkin that actually reach market. I attempt to control for this using the method described in footnote 6.

apparent. The 2000 main season appeared particularly bountiful in most zones, while the 2007 season was highest yielding over all zones. 1997 was a poor year in the eastern part of the country (i.e., Coastal and Eastern Lowlands), but relatively good in the High Potential Maize areas to the west.

Table 6: Mean output value as defined by Liu-Myers yield index (kg/ha)

		1997	2000	2004	2007	2010
Coastal Lowlands	Maize mono	434	1146	649	873	895
	Maize inter	856	1701	949	1892	1253
Eastern Lowlands	Maize mono	521	1407	1289	1094	2611
	Maize inter	711	1762	1352	2047	2489
Western Lowlands	Maize mono	712	720	473	1407	1124
	Maize inter	942	1053	1064	2336	1721
Western Transitional	Maize mono	1250	1979	2272	2038	3253
	Maize inter	1609	2538	2623	3204	3106
High Potential Maize	Maize mono	3655	2551	3554	3335	2297
	Maize inter	3015	3021	3875	3657	2662
Western Highlands	Maize mono	1241	1944	1102	1552	1584
	Maize inter	1654	2118	2067	3156	3311
Central Highlands	Maize mono	1877	2484	1925	2547	2454
	Maize inter	2337	3080	2811	3530	4831
Marginal Rain Shadow	Maize mono	-	1778	593	-	1368
	Maize inter	1060	1709	2124	2760	2068
Total sample	Maize mono	2214	2049	2442	2644	2078
	Maize inter	1934	2338	2471	3063	2789

Note: Maize output is computed using the Liu and Myers index. See text for additional information.

While the Liu and Myers index is useful for standardizing yield across different types of plots, the computation is influenced by both the composition of the plot and local relative prices. In order to recover more consistent estimates of the variables of interest, two control variables are added to the model to deal with potential measurement error. In order to control for the possibly biasing effects of higher valued and higher yielding crops on the same fields as maize, a categorical variable describing the number of crops on a given field (ranging from one to seven, per the above requirements) is included as a control, allowing the intercept of the production function to vary with the number of crops found alongside maize. One additional element of bias is the differences in relative prices between crops. If the relative price ratio between, for example, maize and beans is vastly different across districts, this will produce different levels of yield for the same field composition found in different places. However, because I want to use prices as specific to the household as possible, mimicking the true opportunity cost of production, I use district level prices in the yield calculation but absorb any variation in relative prices via district level dummy variables.

1.2. Fertilizer: nitrogen and phosphorous components

A large number of fertilizer types (both basal and top dressing) are available for purchase and use in Kenya. While diammonium phosphate (DAP) and calcium ammonium nitrate (CAN) are the most commonly used, either as a pair or individually (for more, see Table 8), a number of other types are used by farmers in Kenya. What is most important for yield response is not the type of fertilizer applied, but the amount of constituent nutrients contained within the given inorganic fertilizer. Nitrogen (N), phosphorus (P) and potassium (K) are the three major nutrients most often considered in analysis of fertilizer use. As previously noted, researchers have been concerned for decades about the depletion rate of nitrogen and phosphorus from the soils of SSA (Sanchez et al. 1997). In fact, there is a general lack of evidence on potassium deficiency and maize response to applied potassium in the literature (e.g., Snapp 1998). Not only that, but most fertilizers found and used in Kenya contain mostly nitrogen and phosphorus, with very little or no potassium. For all of these reasons, only the nitrogen and phosphorous components of applied fertilizer types are used in regression analysis.

Inorganic fertilizers contain a regulated ratio of nutrients per unit, making it analytically trivial to separate the amount of fertilizer applied into its constituent nutrient parts. Appendix 2 shows the percent of nitrogen and phosphorous found in one kilogram of each type of fertilizer. With these, I compute the total amount of nitrogen and phosphorous from inorganic fertilizer applied per hectare to maize fields, as shown in Table 7. Note how these values differ from the aggregated fertilizer application rates in Table 4. Using only non-zero fertilizer application values, this table shows that households generally choose fertilizers where the nitrogen component is greater than the phosphorous component, due in part to the presence of top dressing fertilizers. Households in Western Transitional, the High Potential Maize and Central Highlands apply the most fertilizer, on average. Households in the all lowlands zones and Western Highlands are observed applying more fertilizer over time.

Table 7: Mean kilograms of nitrogen and phosphorus applied per hectare (excludes zeros)

		1997	2000	2004	2007	2010
Coastal Lowlands	N	0.6	6.3	0.9	8.2	8.5
	P	0.0	0.0	0.4	8.6	5.6
Eastern Lowlands	N	7.3	11.2	10.4	15.7	25.5
	P	3.5	4.1	4.6	6.0	11.2
Western Lowlands	N	17.1	8.4	8.7	11.0	21.4
	P	7.8	7.6	8.7	9.6	17.6
Western Transitional	N	33.0	32.9	43.1	49.0	46.3
	P	20.6	19.0	20.0	21.2	21.4
High Potential Maize	N	31.3	39.2	42.2	43.8	44.4
	P	26.2	24.8	24.9	23.5	24.1
Western Highlands	N	16.2	17.8	27.9	28.0	41.9
	P	17.6	15.0	19.0	20.2	25.0
Central Highlands	N	36.8	29.4	30.6	30.4	66.7
	P	16.4	14.2	15.7	15.2	16.3
Marginal Rain Shadow	N	-	-	126.4	12.6	22.2
	P	-	-	39.6	10.1	24.8
Total sample	N	29.9	32.0	34.9	37.6	44.5
	P	20.7	19.8	19.7	20.2	21.4

Note: Averages only include observations where fertilizer was used, no matter the sample size.

What I capture in these variables is the amount of nitrogen and phosphorus *applied* to fields in a given season, not the amount of nitrogen and phosphorous *available* in the soil. It is well known, however, that available nutrients and, therefore, soil fertility (or, soil organic matter) are what drive productivity (Bauer and Black 1994). While available nutrients are not observed here, I attempt to control for these features using soil types (i.e., available nutrients vary with soil properties as captured by soil types) and socio-economic variables (i.e., proxies for household level soil fertility and nutrient availability). I describe these in more detail in later sections.

Another important distinction is between the amount of nitrogen and phosphorous *applied* to the field and the amount *absorbed* by the plants. Application of a nutrient does not necessarily imply full absorption, as rates of uptake vary tremendously. Applied nitrogen, for example, is generally used by the plant that season. Phosphorous, however, is a less mobile nutrient and, therefore, the carry over from previous seasons is very important for current season plant growth, making the phosphorous reservoir in the soils an issue of long term investment. Crops generally use only 20 percent of the applied phosphorous in the first year of application (Griffith) which makes modeling yield response to phosphorous difficult without a good understanding of available phosphorous through soil samples or knowledge of past farming practice (see Lanzer and Paris 1981).[8] Furthermore, while it is well-understood that the absorption rate of plants varies with the soil organic matter content and general characteristics of the soil (e.g., Vanlauwe et al. 2000 on savanna soils in West Africa), researchers understand the dynamics of nitrogen far more than those of phosphorous (Merckx et al. 2001) making it difficult to accurately specify which other soil inputs and nutrient factors are driving phosphorous availability and sorption.

Ideally, the researcher would observe households applying different ratios of nitrogen to phosphorous between fields and across time in order to get good estimates of response to applied

---

[8] I cannot control for past farming practice because I do not consistently observe the same field every year (e.g., field one at household one in year one is not necessarily the same as field one at household one in year two). Even at the household level, the gap between survey years means I cannot accurately estimate nutrient carry-over.

nitrogen and phosphorous as separate applied nutrients. However, as previously mentioned, households in Kenya overwhelming choose to fertilize their maize fields under two via two methods (1) DAP only or (2) DAP with CAN in relatively fixed proportions, making the ratio of nitrogen to phosphorous application fairly fixed across households and time. Table 8 shows the percentage of maize fields in a given zone fertilized by type of basal or top dressing. Because of the high degree of correlation between the nutrients, maize response to applied nitrogen and phosphorous application cannot be assessed separately. Given (1) the difficulties in accurately detailing the response to applied phosphorous and (2) the issue of collinearity, this thesis will mostly focus on applied nitrogen while noting both the interaction with phosphorous and the omitted variable bias this method might produce. For more on the collinearity problems with phosphorous, see Appendix 3.

Table 8: Percent of fertilized maize fields with specific basal and top dressing types

	Basal		Top dressing	
	DAP	other	CAN	other
Coastal Lowlands	58	0	47	0
Eastern Lowlands	60	17	71	1
Western Lowlands	89	2	9	9
Western Transitional	92	1	27	40
High Potential Maize	88	15	31	11
Western Highlands	98	1	32	10
Central Highlands	59	31	43	1
Marginal Rain Shadow	100	0	33	0
Total sample	85	12	34	13

Note: Percentages represent portion of all fertilized fields with particular type of fertilizer. Other basal fertilizers include MAP, TSP, SSP, and NPKs. Other top dressing fertilizers include UREA, ASN, and SA. All survey years included.

1.3. Manure and compost

While inorganic fertilizer application is of primary interest here, a considerable number of farmers in Kenya use manure or compost in conjunction with inorganic fertilizers or as a replacement. Non-inorganic fertilizers like fresh manure and compost contain useful macro and micro nutrients. For example, Smaling et al. (1992) use experimental evidence across agro-ecological zones in Kenya to show that manure use can significantly increase yields under certain conditions. Smaling et al. (1992) also find that the interaction between inorganic fertilizer and manure produces favorable results in certain agro-ecological zones. This is likely due to the fact that yield response to applied phosphorous depends on the soil's ability to dissolve phosphorous, which is aided by the presence of acidifying agents. African soils generally have too high of pH values (i.e., above 6.2) to accomplish this task alone, which is why applying manure or compost to the soils is useful (Sanchez et al. 1997). Furthermore, the organic materials in manure and compost add to the carbon-stock of the soil, without which applied nitrogen and phosphorous remain inaccessible to crops.

Table 9: Percent of fields with manure/compost and mean kg/ha applied by users

		1997	2000	2004	2007	2010
Coastal Lowlands	% use	26	27	25	22	41
	Mean kgs	-	-	1047	1076	1374
Eastern Lowlands	% use	59	54	64	54	70
	Mean kgs	-	-	2596	2803	1374
Western Lowlands	% use	13	16	19	32	35
	Mean kgs	-	-	2062	1526	2605
Western Transitional	% use	11	39	25	31	44
	Mean kgs	-	-	1191	1268	1069
High Potential Maize	% use	1	17	17	14	21
	Mean kgs	-	-	1172	1620	1332
Western Highlands	% use	9	14	18	15	25
	Mean kgs	-	-	1629	1497	1935
Central Highlands	% use	14	55	79	85	86
	Mean kgs	-	-	5640	4933	4620
Marginal Rain Shadow	% use	0	89	60	58	53
	Mean kgs	-	-	2864	3189	2077
Total sample	% use	14	29	33	31	41
	Mean kgs	-	-	2919	2694	2337

Note: Mean kilograms calculated from just those who applied (excludes zeros). Actual amounts only collected in last three years of the survey.

While I observe on which fields manure or compost is applied, there is no easy way to verify their nutrient composition without analysis at the household level (Murwira et al. 1995;

Probert et al. 1995), meaning I cannot treat manure application in the same way I do inorganic fertilizer application.[9] However, given evidence of its importance and the observed high use of manure by households in Kenya, I want to ensure those fields with manure or compost applied are accurately distinguished and do so in the production function by including a dummy variable for fields where any amount of manure or compost was applied. Table 9 shows the percentage of fields where manure or compost was applied. For reference, I also include the amount of manure and compost applied per hectare in the last three survey years (i.e., this variable is not collected consistently across years and is, therefore, not used in the production function). These numbers are considerably higher than inorganic fertilizer application amounts, which could either be a function of the lesser nutrient content of manure or the fact that manure and compost are more readily available, particularly on farms that have livestock, and therefore far less costly for households to use.

Not only do manure and compost function as inputs into the maize production system, but they may also serve as a proxy for soil organic matter levels at the field level. Numerous studies in SSA have shown that applying manure to continuously cropped fields slows the rate of soil fertility loss, even when coupled with inorganic fertilizers (e.g., Kapkiyai et al. 1999 in Kenya; Agbenin and Goladi 1997 in Nigeria). Because I do not observe household or field specific soil organic matter levels, manure and compost application rates may serve as a useful proxy for soil quality and available nutrients. Unfortunately I do not observe the same field over time and cannot measure the longer term impacts of applying organic matter to the field. The manure and compost dummy variable used here only captures contemporaneous effects.

---

[9] While this is certainly true, in their survey of the literature, Giller et al. (1997) find that the nitrogen composition in cattle manure in Kenya tends to be much higher than in other SSA countries, likely due to better diets offered to the livestock.

1.4. Legumes

Another non-chemical way to add nitrogen into the soil is to intercrop maize with legumes (e.g., beans, peas, lentils, groundnuts). Leguminous plants have the ability to fix atmospheric nitrogen into the soil and can therefore be important for nutrient cycling and soil development (Groffman et al. 1987; Vansambeek et al. 1986; Ledgard and Stelle 1992) and ecosystem function (Chaplin et al. 1986; Mooney et al. 1987). Rao and Mathuva (2000) used an experimental research station scenario in Kenya to show that intercropping maize and pigeon pea increased maize yields by 24 percent over monocropped maize fields. Similarly, Maobe et al. (2000) find evidence in southwestern Kenya that green manure from leguminous crops increases the profitability of intercropped production systems over that of monocropped systems. Not only do legumes fix nitrogen into soil, but research also suggests that intercropping with pigeon peas increases total phosphorus availability in cropping systems with low phosphorous availability (Ae et al. 1990).

Intercropping with legumes is a feature of a large number of the intercropped fields in Kenya. Given experimental findings and the incidence of legume intercropping in this dataset, I include a dummy variable on fields where maize is intercropped with legumes.[10] Maize intercropped with beans is the most frequent combination, however beans generally fix about half as much nitrogen into the soil than other leguminous crops (Piha and Munns 1987). In fact, Giller and Cadisch (1995) find that beans have such low capacity to fix nitrogen into the soil that they can produce a negative nitrogen balance instead, meaning they use more nitrogen than they fix. For this reason, common beans are not considered a legume in this analysis.

---

[10] In this dataset, the following non-common bean legumes are present: chickpeas, cowpeas, dry peas, French beans, green grams, green peas, groundnuts, njahi (dolichos), njuga mawe (bambara beans), pigeon peas, runner beans, simsim, snow peas, and soy beans.

1.5. Seed rate and type

Seed is an important contributor to yield, both the type of seed used and the quantity. In this dataset, I observe on which fields farmers use new hybrid maize seeds. New hybrid seed is considered yield increasing (Hassan et al. 1998), although recycled hybrid seed is said to have little yield advantage over local non-hybrid. In a paper on technology adoption in Kenya, Suri (2011) focuses on hybrid maize seeds and finds that farmers, in general, are using hybrid seeds when it is profitable for them to do so and to maximize their unobserved comparative advantages. Similarly, there is evidence that local variety seed may perform better on poorer quality soils. For these reasons, I include a dummy variable on fields where new hybrid seeds are used. Table 10 shows the percentage of maize fields farmed with new hybrid seed by year and zone. In general, hybrid seed use is quite high, although mostly skewed towards the higher yielding zones to the west.

Table 10: Percent of maize fields with new hybrid maize seed by year and zone

	1997	2000	2004	2007	2010
Coastal Lowlands	27	25	2	29	40
Eastern Lowlands	25	24	11	41	69
Western Lowlands	16	25	13	30	30
Western Transitional	71	79	71	85	87
High Potential Maize	88	89	91	93	98
Western Highlands	74	82	67	78	89
Central Highlands	87	81	74	81	90
Marginal Rain Shadow	87	67	38	51	91
Total sample	66	66	59	72	79

The seed rate, calculated as kilograms of maize seed applied per hectare, is computed equivalently across monocropped and intercropped fields. Seed rate (specific to maize) likely varies with the number of crops on the field, making it a useful control variable for fields where area planted to maize was relatively smaller. Table 11 shows how the average seed rate has varied over time and by the number of crops per field. Seed rates are surprisingly similar by number of crops on the field; even maize fields with six other crops have seed rates similar to monocropped fields, meaning the criterion used here to capture maize fields is doing an adequate job of picking out fields where maize is the dominant crop in terms of both planted area and potential revenue. Nevertheless, these values are averages, so seed rate per field is used as a control in the production function.

Hybrid seed is also said to have an even larger positive impact when coupled with inorganic fertilizers (Ellis 1992). Unfortunately, however, I am unable to estimate the interaction between hybrid seed use and inorganic fertilizer application because farmers in Kenya almost always use the two in tandem. About 53 percent of households in our data set always use hybrid seeds on their maize fields, of which 89 percent always use fertilizer.

Table 11: Mean maize seed rate (kilograms per hectare) by number of crops on field

	1997	2000	2004	2007	2010
1 crop	23.5	22.6	22.2	23.4	24.5
2 crops	23.2	21.1	21.1	23.1	24.7
3 crops	20.2	22.6	21.0	22.0	23.6
4 crops	16.9	20.9	20.3	20.7	22.8
5 crops	19.5	21.8	19.9	19.5	21.9
6 crops	20.4	22.3	17.5	20.9	22.2
7 crops	20.0	20.1	20.7	20.0	21.2
overall	22.3	21.7	20.9	22.1	23.4

1.6. Field and farm size

Because most continuous variables included in the production function are in "per hectare" format, I also include a variable to control for total hectares on a given maize field. One might assume that on smaller fields farmers might attempt to fit a larger number of crops or pack rows more tightly while devoting larger fields to maize monocropping. This variable should absorb any remaining variation attributed to the size of the field. Table 12 shows the average and standard deviation in field sizes. In generally, fields are small, particularly in the highland areas where population density is highest. The High Potential Maize Zone is the only one where field sizes are consistently above one hectare on average.

In this dataset, the size of individual maize fields is also highly correlated with the total farm size (correlation coefficient of 0.6). For many years, researchers have observed an inverse relationship between farm size and productivity (Chayanov 1962; Sen 1962; Berry and Cline 1979). A large number of possible explanations are provided, including (1) labor market dualism (i.e., households with smaller farms have a lower opportunity cost of labor and, therefore, apply less labor to the farm than larger farms when equating the time spent on the farm with the marginal value product of labor), (2) the availability of decreasing returns to scale technology, (3) market failures limiting the amount of inputs available for larger farms, and (4) village-specific effects that cause substantial differences in prices, soil and wages (see Barrett 1996 for a review of the literature). To the extent that field and farm size are related at the household level, this variable will also provide insight into the farm size and maize productivity phenomenon.

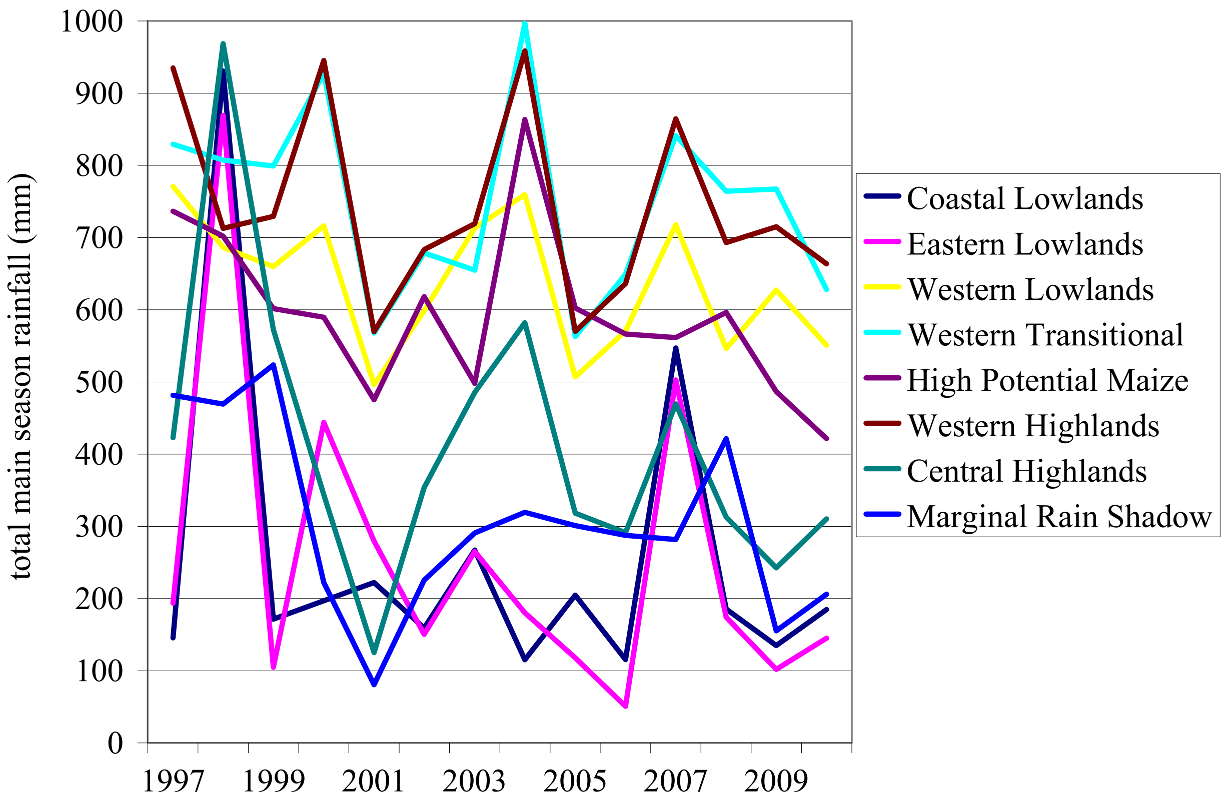
Table 12: Mean and standard deviation of maize field size (hectares)

	1997	2000	2004	2007	2010
Coastal Lowlands	0.75	0.75	0.92	0.76	0.67
	(0.76)	(1.2)	(0.95)	(0.70)	(0.83)
Eastern Lowlands	0.86	0.60	0.65	0.67	0.50
	(0.74)	(0.61)	(0.50)	(0.64)	(0.35)
Western Lowlands	0.46	0.45	0.45	0.34	0.38
	(0.34)	(0.32)	(0.30)	(0.18)	(0.23)
Western Transitional	0.58	0.62	0.49	0.48	0.43
	(0.43)	(0.54)	(0.39)	(0.40)	(0.37)
High Potential Maize	0.99	1.0	0.81	0.90	0.86
	(0.88)	(0.88)	(0.74)	(0.83)	(0.83)
Western Highlands	0.36	0.38	0.29	0.25	0.27
	(0.24)	(0.34)	(0.24)	(0.18)	(0.22)
Central Highlands	0.36	0.28	0.25	0.22	0.21
	(0.26)	(0.30)	(0.17)	(0.15)	(0.15)
Marginal Rain Shadow	0.51	0.26	0.43	0.36	0.40
	(0.26)	(0.14)	(0.23)	(0.20)	(0.19)
Total sample	0.66	0.67	0.56	0.57	0.53
	(0.66)	(0.73)	(0.56)	(0.63)	(0.59)

1.7. Rainfall

In Kenya, drought and erratic rainfall levels are characteristic of the production system. Using rain station data from the National Weather Service Climate Prediction Center (CPC) as a part of their Famine Early Warning System (FEWS) project, rainfall conditions are estimated at the village level using GPS coordinates taken during data collection for use in this analysis. Figure 6 shows average main season rainfall levels across the agro-ecological zones. Notice how rainfall amounts can vary significantly over time and across space, even over a relatively short period of time.

Figure 6: Average main season rainfall across agro-ecological zones over time



Contemporaneous rainfall is essential to the production function given its role in converting chosen inputs into yield. In Kenya, two characteristics of actual rainfall are important: (1) total rainfall over the agricultural season of interest and (2) the distribution of rainfall over that season. In this dataset, and perhaps generally speaking, total rainfall and rainfall stress are highly correlated (correlation coefficient of 0.86), meaning the two should not be included together in the production function. Instead, I use rainfall stress given the importance of fairly continuous rainfall amounts in maize production. Here, rainfall stress is measured as the percentage of days in a 20 day period during growing season where the rainfall level dips below 40 millimeters.

Table 13: Average total main season rainfall and rainfall stress by zone and year

		1997	2000	2004	2007	2010
Coastal Lowlands	Total rain	158	225	120	540	205
	Rain stress	0.66	0.47	0.80	0.29	0.71
Eastern Lowlands	Total rain	195	452	183	492	152
	Rain stress	0.64	0.33	0.60	0.16	0.62
Western Lowlands	Total rain	771	709	766	718	565
	Rain stress	0.09	0.05	0.06	0.22	0.34
Western Transitional	Total rain	828	928	989	840	628
	Rain stress	0	0	0	0.14	0.12
High Potential Maize	Total rain	742	595	862	559	456
	Rain stress	0.16	0.29	0.13	0.34	0.46
Western Highlands	Total rain	932	943	955	856	683
	Rain stress	0	0.11	0.05	0.22	0.17
Central Highlands	Total rain	423	383	562	491	300
	Rain stress	0.46	0.45	0.46	0.39	0.50
Marginal Rain Shadow	Total rain	495	217	317	284	182
	Rain stress	0.16	0.74	0.57	0.63	0.81

Note: Total rainfall is observed in millimeters. Rainfall stress is the fraction of days in a 20-day period during the growing season where rainfall level dipped below 40 mm.

Table 13 shows the average total main season rainfall and rainfall stress by zone. In general, maize performs best under 500 to 800 millimeters of rainfall per season (Ovuka and Lindqvist 2000). While some zones always fall within this range (e.g., Western Lowlands), others are consistently below (e.g., Eastern Lowlands), and others with averages mostly above (e.g., Western Highlands). Rainfall levels below 250 millimeters over a main season are considered inhospitable to maize production (Kironchi et al. 2006), however many households in this data set reported maize yields despite low levels of main season rainfall, meaning farmers and/or local variety seeds have adapted to drought-like conditions. On the other hand, rainfall levels above 1000 millimeters approach flood conditions and may simply wash away fields and plants (e.g., 1998 El Nino year in Eastern Kenya). Again, there are several districts where rainfall levels are reported over 1000 millimeter during one of the survey years included in this analysis and where maize is still harvested.

1.8. Soil

The health and type of soil used for production is important for both plant growth, how the soils take in applied inputs (e.g., fertilizer), and output. Soil scientists recognize the heterogeneity in soil conditions across space and time and often attempt to understand this diversity by taking and analyzing soil samples at the field level to test for available nutrients and soil organic matter status. While some of these efforts have been large-scale (e.g., Snapp et al. 2010 in Malawi), most studies using field level soil data are limited to a small geographic range (e.g., Marenya and Barrett 2009b in Kenya) making extrapolation of conclusions to a larger-level more difficult. What these careful pieces of analysis do tell us, though, is that soil conditions are critical variables in understanding and estimating yield response.

For this study, data on time invariant soil characteristics (i.e., drainage, depth, texture) and FAO soil classifications are available at the village level from the Kenya Soil Survey and the Ministry of Agriculture from data originally collected in 1980 (see Appendix 12 for a map). Soil composition, though, can vary dramatically from one plot to the next, meaning the use of village level averages of clay and sand content fail to accurately characterize each field level observation. Instead, the FAO soil classification system, which groups soils based on their formation process and overarching properties, is a better village level indicator of soil type. In this dataset, ten different types of soils are included (i.e., Cambisols, Ferralsols, Phaeozems, Luvisols, Greyzems, Podzols, Regosols, Solonetz, Rankers, Vertisols) with a large range in the number of villages falling into each category. In order to understand the inherent productivity levels of these soils, all else equal, I include each FAO classification as a dummy variable (intercept) in the model.

Because these soil types do not completely capture the variability in soil fertility conditions at either the household or field level which are most often correlated with past farm management and soil fertility decisions. Instead, I use various proxies for household and field level soil fertility to attempt to control for whatever important variability remains, some of which have already been mentioned: (1) applied manure and compost may act as a proxy for soil organic matter levels given the high levels of micronutrients available in both, (2) households may use local variety seeds on fields they presume to have poor soils and better hybrid seeds on fields they perceive to have fertile soils, and (3) household level socio-economic characteristics are likely correlated with farm management decisions and soil health (see next section).

1.9. Socio-economic variables as proxies

There are a number of unobserved characteristics of the production setting important for analysis which can be accounted for using socio-economic variables as proxies. For example, while I observe soil characteristics at the village level, there is good reason to believe that soil organic matter and characteristics vary by household in Kenya. Tittonell et al. (2005) grouped farms in western Kenya by socio-economic status and found that soil fertility and nutrient flows varied considerably between groups, concluding that both inherent soil properties and management explain variability in soil fertility status. Similarly, Marenya and Barrett (2009b) find that the more degraded plots are cultivated by the poorest households in the villages they study in western Kenya. With evidence that the socio-economic status of a household is correlated with soil fertility at the farm level, I include a measure of asset wealth[11] in the production function as a proxy for soil quality. However, because the Mundlak-Chamberlain device uses variation within the household to estimation parameters, I use a five-year average of total asset wealth of the household as an indicator of its longer term socio-economic standing so as to compare across households instead of within.

2. Pooling and Grouping Methods for Estimation

Maize growing conditions in Kenya can vary wildly from place to place and, therefore, so too can the input technologies required or useful for production. This known heterogeneity in conditions means that all sampled households do not necessarily encounter the same response to all inputs, particularly fertilizer, the focus of this analysis. Grouping households by similarities in geographic and environmental conditions allows for more targeted estimates of the response to fertilizer and consideration of the specifics of the production setting to which households are privy. A balance must be struck between too few and too many groups, however, because a small sample size will limit degrees of freedom and the variation necessary to estimate parameters with confidence (i.e., small sample bias). As such, I condition the coefficients on nitrogen response on (1) where geographically the field is located (i.e., by zone group) and (2) on what type of soil (i.e., by soil group). I test the hypothesis that the groups described below are necessary using a Chow Test and report the results in the next section.

[11] Asset wealth is defined here as the total value of livestock, farm equipment and large household objects consistently recorded across all years of the survey.

The first important grouping is by agro-ecological zone. Conditions in the Coastal Lowlands and Marginal Rain Shadow zones are very different from the others, with very low rainfall levels and far less incidence of maize production; for these reasons, maize fields in these zones are dropped from analysis. Within these six remaining zones, however, response to fertilizer application likely differs given heterogeneity in the production environment available to the households. In order to produce nitrogen response conditional on observed agro-ecological characteristics, I group households into groups with similar environmental features. Because relatively large sized groups are necessary to get good estimates of conditional response, I aggregate the six remaining agro-ecological zones into three groups: (1) lowlands, consisting of the Eastern and Western Lowlands Zones (same elevation, less rainfall, lower rates of fertilizer use); (2) high potential and transitional areas, consisting of the High Potential Maize and Western Transitional Zones (same districts, high rates of fertilizer use); and (3) highlands, consisting of the Western and Central Highlands Zones (same elevation, similar rates of fertilizer use). Each of these groups has well over 100 households (see Table A.3 in Appendix 5 for standard deviations of production function variables at this level) within which to estimate how relative homogeneity in conditions contributes to differences in fertilizer response.

Within and between these zone groups, similarities in soil type may also contribute to differences in fertilizer response. I further consolidate the ten FAO soil types used as intercepts into six groups producing a larger number of households in each group (i.e., over 100) to facilitate adequate variation. Grouping to this level ensures that individual soil categories are not influenced by the other local agro-ecological features found alongside a given soil, particularly when a given soil is only found in one area of the country. FAO soil types were grouped using their definitions and help from Table 1 of (IUSS Working Group WRB 2007) where the soil formation process is described. Table 14 shows how soils are grouped based on their similar soil characteristics. This grouping scheme represents the end of a long process of trial and error in grouping soils through various methods using available data. For more on this process and how this final grouping system evolved, see Appendix 4.

Table 14: Characteristics of soil groups

Soil group number and criteria (number of villages)	Number of villages by soil classification	Number of villages by agro-ecological zone
1 Volcanic landform: Regosols and some Podzols (25)	Podzols: 2 Regosols: 23	High Potential Maize Zone: 9 Central Highlands: 16
2 High humus or highly productive: Phaeozems, Luvisols, Greyzems, Cambisols (21)	Cambisols: 4 Phaeozems: 6 Luvisols, 10 Greyzems: 1	Eastern Lowlands: 1 Western Transitional: 1 High Potential Maize Zone: 11 Western Highlands: 2 Central Highlands: 3 Marginal Rain Shadow: 3
3 Rankers with more sand (25)	Rankers: 25	Coastal Lowlands: 4 Eastern Lowlands: 11 Western Lowlands: 2 Western Transitional: 4 High Potential Maize Zone: 1 Western Highlands: 3
4 Rankers with less sand (20)	Rankers: 20	Western Lowlands: 1 Western Transitional: 6 High Potential Maize Zone: 7 Western Highlands: 5 Central Highlands: 1
5 Vertisols, Ferralsols, and Podzols with high clay and inadequate drainage (9)	Ferralsols: 1 Podzols: 7 Vertisols: 1	Eastern Lowlands: 1 Western Lowlands: 7 Marginal Rain Shadow: 1
6 Very shallow or very poorly drained soils found in swamps, reefs or erosional plains (5)	Podzols: 3 Solonetz: 2	Coastal Lowlands: 3 Western Lowlands: 2

Note: Rankers are split into two groups based on village-averaged sand composition; group three contains villages with more than thirty percent sand while group four has villages with less than thirty percent sand. The two grayed groups represent conditions inhospitable to maize growth and/or fertilizer response. These villages are excluded from the production function estimation.

Group one contains villages with a volcanic landform, another variable collected in the soil survey. Technically speaking, true volcanic soils are referred to as Andosols in the FAO classification system. While no Andosols appear in this dataset, a number of villages with Regosols and Podzols are found in volcanic areas, meaning the soils were likely influenced by past volcanic activity and should be highly fertile. Group two contains the FAO soil types defined based on their high humus (i.e., the nutrient rich material resulting from the decomposition of organic matter [Smillie and Gershuny 1999]) levels (Phaozems, Greyzems and Luvisols) which were originally under forest or grassland and should contain high levels of organic matter in top soils. The highly productive Cambisols are also included in this group given similarities in fertilizer response and general agricultural productivity. Groups three and four contain all of the Rankers, by far the most frequently occurring of the soil types. Rankers are generally shallow, found over rocks and, therefore, have low water-holding capacity. However, due to the very high number of villages with Ranker-classified soil, I split this group into two by village-averaged sand composition level; group three contains all villages with more than thirty percent sand, while group four contains those with less than thirty percent sand.

The remaining two soil groups (five and six, as grayed in Table 14) are left out of analysis due to insufficient conditions for maize growth and/or fertilizer use and, expectedly, the very low use of fertilizer in these areas. Moreover, the fact that very few households in these areas apply fertilizer makes it impossible estimate the yield response to applied nitrogen and phosphorous. Group five contains all villages with Vertisols ("black cotton soil"), Ferralsols (deeply weathered with low water storage capacity), and the remaining Podzols with high clay and inadequate drainage. Group six contains very shallow or very poorly drained soils that are found in areas with coral reef, plains formed by heavy erosion, or swamps.

3. Model Specification and Testing

In this section, I describe the production function used in estimation and report the results of a range of regression diagnostics and of Chow Tests for the grouping methods described in the previous section. A modified quadratic function is estimated with care to include interactions that have conceptual or theoretical meaning instead of all possible interactions dictated by a true quadratic model. The following model serves as the basis for testing (variable name abbreviations are described in Table 5).

$$Y_{ijt} = \alpha_1 + \beta_1 N_{ijt}{*}zone + \beta_2 N_{ijt}^2{*}zone + \beta_3 N_{ijt}{*}P_{ijt}{*}zone + \beta_4 N_{ijt}{*}soil + \beta_5 N_{ijt}{*}rain_{jt}{*}zone + \beta_6 seed_{ijt} + \beta_7 seed_{ijt}^2 + \beta_8 hect_{ijt} + \beta_9 hect_{ijt}^2 + \beta_{10} asset_j + \beta_{11} asset_j^2 + \beta_{12} rain_{jt} + \beta_{13} manure_{ijt} + \beta_{14} hybrid_{ijt} + \beta_{15} legume_{ijt} + \beta_{16} crop_{ijt} + \beta_{17} dist + \beta_{18} FAO + \beta_{19} year + c_j + \mu_{ijt} \quad (19)$$

Several variables vary by field, household and year (i.e., nitrogen, phosphorous, seed rate, hectares, crops per field, manure, hybrid, legume); others vary by household and year (i.e., rain stress); the remaining are constant characteristics of the household over time (i.e., zone, dist, FAO, soil). Then, as an added control, I include a year variable to absorb all remaining variation in yield attributed to those time-specific characteristics not already accounted for through other observable variables. Finally, the c term represents the Mundlak-Chamberlain device.

Before presenting the results of this regression, I run several diagnostic tests on my model to ensure its appropriateness. Functional form is evaluated by inspecting residual plots. A scatter plot of the residuals against predicted values confirms that the functional form is appropriate. I investigate the variation in and multicollinearity between parameters. For the former, I show that standard deviations across all variables are high (Table A.3 in Appendix 5), meaning variation should be sufficient for estimation. Multicollinearity is rejected by estimating a condition score of 23.27, which is under the 30 threshold suggested by Belsley et al. (1980), meaning I should be able to parse out the effects of individual parameters. I test the assumption of homoskedasticity using a Breusch-Pagan test. Results (F-statistic of 5.62, well within the critical range at the 99 percent confidence level) show that I should reject the null hypothesis of homoskedasticity and, moreover, that the standard errors computed by OLS are unbiased. While this pattern is expected in survey data coming from a large sample, I account for non-constant variance by computing robust standard errors clustered at the household level, a common solution to heteroskedasticity (Wooldridge 2009b). Clustering at the household level has the added benefit of making standard errors robust to serial correlation, meaning they are able to control for how input decisions by a household in one year affect response in subsequent years.

I also want to test whether or not it is appropriate to include the zone and soil group interactions on nitrogen, allowing the response of nitrogen to vary across space. The first line of equation 19 includes these conditioning methods, as described in Section 2 of this chapter. By dropping the zone and soil group interactions from the nitrogen and phosphorous terms and comparing it to the full model in equation 19, I can test the null hypothesis that these interactions should not be included or, said differently, that the response to nitrogen and phosphorous does not vary by zone and soil group. A Wald test yields an F-statistic of 3.36, which is significant at the 99 percent confidence level; therefore, I can reject the null hypothesis that the fertilizer response variables are the same across Kenya. The model as described in equation 19 is, then, the one used in the regression results that follow.

I also investigate the need for controlling for unobserved household heterogeneity through the Mundlak-Chamberlain device. This is done by comparing the results of the model with and without the means of time-variant variables through a Wald test, which is essentially a comparison of the pooled OLS estimation versus CRE. The comparison yields an F-statistic of 4.56, which is significant at the 99 percent confidence level, meaning I can reject the null hypothesis that unobserved household heterogeneity is not present. Furthermore, in order to use the Mundlak-Chamberlain device to control for unobserved household heterogeneity, there must be sufficient variation in the variables both within (i.e., at the household level over time) and between (i.e., between households in a given time period). For all variables in the production function, Table A.3 in Appendix 5 shows sources of variation. This table shows relatively high standard deviations across all variables both at the household level and between households, no matter what sample of households is used. The CRE model with Mundlak-Chamberlain device, therefore, is used to control for the unobserved household heterogeneity and produces unbiased estimates of the parameters of interest.

4. Production Function Estimation Results

In this section, I present the results of the yield response model in both raw regression format and, because raw regression results can be difficult to interpret when individual variables appear multiple times in the same model, as computed average marginal effects (AMEs).[12] I relate findings back to hypotheses of many of the included terms then, specifically, to applied fertilizer. Marginal products and average products for nitrogen are calculated at the level of disaggregation used in the household level profitability analysis that follows in the next chapter.

4.1. Regression results and marginal effects

The production function estimation results (Table A.5) and average marginal effects of all explanatory variables (Table A.6) can be found in Appendix 6. In both, I compare the results

[12] This is as opposed to the marginal effects at the average (MAEs), which is the computed marginal effect of the average individual in the specified group. However, I did calculate these and found them to be similar for some areas, but mostly not estimable when there were too many zero values of inputs.

of the same model estimated under pooled OLS, fixed effects (FE), and correlated random effects (CRE). Production function coefficient estimates are fairly similar across the three, providing confidence into the structure of the model. Some coefficients, however, are most similar between the fixed effects and correlated random effects models (for instance, the coefficients on legume and rain stress), showing how controlling for within-household variation leads to different estimates. Notice, too, that several of the variables in the Mundlak-Chamberlain device are statistically significant, another indicator of the importance of controlling for the consequences of unobserved household level heterogeneity. For most of what remains, I will refer to the correlated random effects (CRE) estimates.

The signs and significance level of most of the squared terms provide further validity to using a quadratic functional form. Most of the squared terms generate negative and statistically significant estimates, meaning a diminishing marginal returns relationship is appropriate for many inputs, including nitrogen, seed rate, and asset wealth. This means that for most inputs that are included as continuous variables, they contribute to yield positively up to a certain point (although with each additional unit adding less to yield than the one before it) after which negative returns set in.

As discussed in Section 2 of this chapter, the nitrogen variable is interacted with several variables to create response conditional on agro-ecological surroundings. The lowlands are the areas with the most concave and steepest slope on nitrogen, with the highlands and higher potential areas having less concavity. Furthermore, the lowlands areas have a much higher response to combined nitrogen and phosphorous (the coefficient on the interaction is about 1.4) than the other two areas (with coefficients around 0.3). Not only does this coefficient pick up on the differences in response to combined nitrogen and phosphorous, but also the differences in the ratio of applied phosphorous to applied nitrogen across space. In the eastern lowlands, for example, households are more likely to use top dressing with basal (which would lower the phosphorous to nitrogen ratio) whereas households in the highlands and higher potential areas are more likely to apply only basal (which would increase the phosphorous to nitrogen ratio). In some sense, this interaction term is acting as a control for the type of fertilizer used.

Once nitrogen response is conditioned on geographic zone variables, the interactions with soil groups do not produce statistically significant coefficients. There are a number of reasons this might be the case. First, the individual FAO soil classifications are already included intercepts, which should show the overall productivity of different soil types once all inputs and other environmental factors are held constant. It could be the case that while these soils have different inherent productivity levels, their responses to fertilizer are not very different between soil types. Second, I lump all soils into four different categories, which could be far too high level of aggregation to tease out how soil characteristics contribute to differences in fertilizer response. This, however, was done in order to have enough observations and variability across space to provide more accurate information on how soil quality contributes to fertilizer response. Third, perhaps soil formation properties (off of which the FAO classifications are based) are not as important to fertilizer response as the actual nutrient composition of the soil (see Marenya and Barrett 2009a), for which we do not have data.

The final interaction with nitrogen is rainfall stress. One would hypothesize that in areas with high levels of rainfall stress (correlated with low levels of rainfall) would have a lower response to fertilizer than areas with less rainfall stress and higher levels of total rainfall. Moreover, when simply interacting nitrogen with rainfall stress, there was no significant relationship which, in the fixed effects and correlated random effects models, means that within-household variation in rainfall conditions does not change the response to fertilizer significantly. Instead, I conditioned the interaction between nitrogen and rainfall stress on the zone groups, which produced statistically significant results. The coefficients on the lowlands and highlands interactions are positive, which might seem counterintuitive, but is actually showing the differences in districts included in those groups. For example, in the lowlands group, the Eastern Lowlands have more rain stress than the Western Lowlands but also more fertilizer use and, perhaps, higher fertilizer response. Similarly, in the highlands, the Central Highlands have more rain stress than the Western Highlands, but also more fertilizer use. The sign on the high potential group is negative, more in line with expectations, which is a product of relative similarly in rain stress conditions across this group and, therefore, a closer approximation of how lower levels of rainfall contribute to lower returns on fertilizer application.

Several other inputs exhibit the expected signs. As hypothesized, applied manure and compost contribute positively to maize yield, either as a contemporaneous input or as a proxy for the soil organic matter level of the field. When I tried interacting the manure dummy variable with other inputs, particularly nitrogen, the coefficients never turned out significant. This could be due to the measurement error when combining all applied manure and compost into a single dummy variable or the fact that I do not observe longer term applied organic matter trends. Either way, it is important to note that even after controlling for applied fertilizer, having manure or compost on the field does contribute positively to maize yield. This trend is also true of new hybrid seed. All else equal, using new hybrid maize seeds leads to higher maize yields. As mentioned previously, I did not interact hybrid with nitrogen given the very high level of collinearity between fertilizer users and hybrid users, making the combined response very difficult to isolate. I did, however, interact the hybrid dummy variable with rain stress. As hypothesized, this term is negative (although not significant in the correlated random effects model), meaning that hybrid seeds are not necessarily a useful choice for households in lower rainfall environments. Having a legume intercropped on a maize field, another potential contributor to available nitrogen in the soil, produces a negative coefficient in my model, although not significant in the correlated random effects model. There are several reasons for this unexpected sign. First, my dependent variable is not pure maize yield but is, instead, calculated using the Liu and Myers index where output is a function of the relative market value of crops. This means that I might not be able to isolate the actual maize output gain given the output of the leguminous crop is also included. Second, I also control for the number of crops on the field, which may be absorbing the legume effect given the high number of legume intercrops in this data set. And, third, the positive gains to intercropping with legume might be longer term or more pronounced in the next season. Given I cannot track the same field over time, I am only able to measure how legume intercrops affect maize yield in a same season.

In terms of other biophysical relationships, the coefficient on rainfall stress is negative and significant, meaning the more intermittent the rainfall, the less maize will be produced, as expected. A quadratic term is not included here because it did not fit the data; a negative linear relationship was much more appropriate. The different soil type dummy variables (one per FAO soil classification) yield significant and mostly expected results. All else equal, Greyzems, found over once-forested landscape with high soil organic matter, have the highest maize yield. Regosols, found in volcanic landscapes in this data set, are the next most productive, followed by Rankers, the most frequently occurring, and Podzols, the other soil type found in volcanic areas.

The two variables that seek to capture the size and sophistication of production—hectares of the field and asset wealth per hectare—produce telling results. The hectares variable exhibits a positive and increasing relationship, as juxtaposed to the other quadratic terms in the model. This means that larger fields, all else equal, are more productive. Given the high degree of correlation between field size and farm size, this suggests that larger smallholder farms are more productive, keeping in mind that this relationship is found for maize fields in the range of 0.06 to 7 hectares only. Conclusions about farm size efficiency cannot be inferred for fields over 7 hectares. However, given most variables are in per hectare format, this variable might simply be controlling for whatever variation in per hectare variation remains at the field level. The asset variable, also measured per hectare, is included to measure how the wealth level of households translates into productivity. This could be interpreted as how much capital they can access or, as suggested in the literature, how well they are able to care for the fertility of their land. Either way, I observe a positive but diminishing relationship, meaning more asset rich households (per hectare) do have a yield advantage, but only up to a point.

While the control variables do not necessarily have important economic interpretation, they do provide insight into their importance as controls. First, the seed rate variable is significant and, therefore, should be controlling for some of the differences in field type or, more importantly, how much of the field is occupied by maize. Second, for what the seed rate does not control, the crops per field dummy variables appear to be doing an excellent job. These seven variables exhibit an astonishing step-like pattern, each one increasing over the one before it. This shows that the dependent variable is increasing in the number of crops per field, and that this variable does control for that relationship. Third, the district dummy variables are included to pull out whatever spatial variation remains not already accounted for in the zone interactions, soil, or rain stress variables. In only a few cases are these variables significant, meaning those observable characteristics are mostly picking up on what makes maize yield response in one district different from the next. And, finally, the year dummy variables control for whatever other temporal variation remains that cannot be accounted for in observed rainfall patterns. Given these are mostly significant, we know that something other than rainfall drives differences over time. However, as previously mentioned, spatial and temporal variation in relative prices could be absorbed by these variables, meaning their interpretation are not necessarily clean.

4.2. Marginal and average product of nitrogen

Given specific interest in yield response to applied fertilizer, this section will look at the marginal physical product (MPP) and average physical product (APP) of nitrogen. As shown in Table A.6, the overall marginal effect of nitrogen is 16.65 in the CRE model, meaning a one kilogram per hectare increase in the amount of applied nitrogen will increase maize yield by about 16.65 kilograms per hectare, all else equal. This value is similar to other overall, highly aggregated marginal products of nitrogen found in the literature throughout SSA. For example, Yanggen et al. (1998) find an average maize response to nitrogen of 17 from studies across all of Eastern and Southern Africa. What I am interested in, though, is local level marginal and average products where variation across space is considered. As such, I calculate these values by district and soil group, where the variation comes from differences in zone, soil group, rain stress levels, and ratio of past phosphorous to nitrogen application.

Also recall that farmers make decisions about input use at planting with uncertainty about how the season will unfold. The main input into the production process that is uncontrollable by the farmer and not known at the start of the season is rainfall. *Expected* rainfall conditions, then, guide farmers' choices about what to plant and with what combination of inputs prior to the main agricultural season. While pre-planting rainfall expectations do not necessarily have their place in the production function (contemporaneous rainfall was used there), they do enter the marginal and average product calculations when modeling farmers' perceptions of output responsiveness to fertilizer application. Technically, the Meteorological Department in Kenya releases its weather forecast information to extension agents and farmers sometime in March which details the prospects for long rains and advises farmers to liaise with the Ministry of Agriculture on the timing of farm operations as well as the crops that would do well under the anticipated weather patterns. However, the extent to which farmers have access to and use this information is not known, meaning using government forecasts as a proxy for farmers' rainfall expectations is not a good strategy here. Instead, I use a six-year moving average of past rain stress levels as a measure of expected rainfall conditions in the coming main season. A six-year average means that no one year has too large of an effect on the average but that recent conditions are taken into account.

Table 15 shows these calculated marginal and average products of nitrogen by district and soil group. This table can be compared to the district and soil group averages of certain variables in the production function found in Table A.4. What this table does not include is the standard deviations in marginal and average products, which are considerably high across the board. For example, the sample-averaged marginal product has a coefficient of variation of about 70 percent while the average product has a coefficient of variation of about 50 percent. This, too, is an important finding; while I am able to capture some of the things that make maize response to fertilizer different across space and time, I am unable to model the subtleties across households and fields. This means that while I can estimate average fertilizer profitability over a district and soil group, it is important to consider what makes response unique at the household level when making targeted prescriptions on fertilizer use. Note that the marginal and average products are very similar for any one place given the chosen functional form; the first derivative of a quadratic function with respect to a particular input (i.e., the marginal physical product) is essentially equivalent to the production function divided by the input (i.e., the average physical product). The similarities in calculated values should, then, not come as much of a surprise.

Table 15: MPP versus APP of nitrogen by district, soil group and year

Province	District	Soil group	MPP					APP				
			1997	2000	2004	2007	2010	1997	2000	2004	2007	2010
Eastern	Machakos	3	43.7	43.4	34.4	40.7	40.8	42.8	43.3	39.9	46.4	54.1
	Makueni	3	37.7	39.1	35.4	38.6	28.5	39.3	40.6	40.8	45.2	41.8
	Meru	1	16.8	18.1	18.2	18.7	17.7	18.8	20.1	20.4	21.0	20.3
	Mwingi	2	47.6	48.0	46.8	52.9	47.5	47.1	47.9	64.8	63.7	49.2
	Mwingi	3	41.0	41.4	39.5	45.5	45.6	42.4	50.4	41.5	53.3	63.1
Nyanza	Kisii	2	19.5	18.7	16.2	18.6	17.6	21.6	20.2	19.4	21.0	21.0
	Kisii	4	17.2	16.2	15.5	16.0	15.8	18.7	17.7	17.5	18.3	19.4
	Siaya	3	29.5	27.1	30.1	32.4	26.6	-	-	35.8	37.6	35.7
	Siaya	4	33.4	31.2	33.3	36.2	25.0	34.0	43.4	42.1	45.4	37.0
Western	Bungoma	2	19.0	18.5	18.6	16.4	16.6	21.1	21.6	21.8	21.1	20.4
	Bungoma	3	9.5	9.3	7.9	9.5	9.0	12.8	13.0	12.5	13.4	13.1
	Bungoma	4	15.7	16.6	14.3	13.2	11.3	18.5	19.9	18.7	18.0	16.2
	Kakamega	2	15.8	13.7	11.9	14.6	12.1	20.0	19.6	18.5	19.7	18.4
	Kakamega	3	11.7	11.5	10.1	9.0	9.4	14.5	13.8	13.4	13.1	13.7
	Kakamega	4	15.3	16.0	14.6	14.2	15.3	16.0	17.4	16.1	16.1	16.8
	Vihiga	3	11.1	9.5	8.5	9.1	8.7	12.3	10.7	10.5	11.4	11.4
	Vihiga	4	15.5	13.2	13.2	13.9	13.5	17.1	14.8	15.4	16.0	16.4
Central	Muranga	1	22.1	19.5	19.4	21.2	19.4	25.3	22.2	21.2	22.5	22.3
	Muranga	4	26.4	23.8	22.7	24.5	23.4	28.1	25.9	23.8	26.1	24.5
	Nyeri	1	21.3	18.3	17.1	20.0	18.9	23.5	20.6	20.3	22.2	21.5
	Nyeri	2	27.4	25.3	24.7	25.8	25.0	29.6	24.9	26.4	27.4	26.7
Rift Valley	Bomet	1	14.7	14.4	14.7	14.9	13.9	17.2	16.2	16.7	16.7	16.0
	Nakuru	1	11.0	10.6	7.6	8.7	8.2	13.1	12.9	9.8	10.9	11.4
	Nakuru	2	13.2	14.4	12.1	10.4	10.9	15.3	16.2	14.3	12.6	13.5
	Nakuru	3	13.5	13.5	11.6	11.4	11.3	15.6	15.5	13.6	13.1	13.7
	Narok	1	5.7	7.9	6.0	4.5	4.7	7.8	9.7	8.2	5.8	6.6
	Trans Nz.	4	14.0	10.9	10.8	10.3	11.1	17.5	15.6	15.6	15.4	15.8
	Uasin Gis.	1	10.9	10.7	9.0	8.0	8.6	13.6	13.8	12.3	11.5	12.2
	Uasin Gis.	2	15.0	12.8	12.3	9.8	10.3	17.9	17.4	17.0	15.6	15.5

Note: Refer to equations 4 and 6. Values are computed at the field level, then averaged to the district and soil group level. These values are used in the profitability analysis that follows.

What it surprising are the vast differences in responsiveness estimated across the country. I find differences in marginal product ranging from 5 to 50 with average products approximately the same. Given I observe large differences in the amount of fertilizer applied, maize yield, and conditions of production, the differences in maize responsiveness are likely not as extreme as they appear at first glance. Furthermore, when testing the model and separately estimating the production function and, therefore, marginal products for the lowlands versus the other areas, estimates were very similar, meaning the fact that the two groups are estimated together is not leading to vastly different values. In general, I find higher marginal and average products in the lowlands areas where fertilizer has only more recently been a feature of maize production. Then, in the areas where farmers have been using fertilizer in large amounts for a much longer period of time, the marginal and average products are much lower. While this might seem counter-intuitive, there is evidence that some areas of western Kenya have over-used nitrogen fertilizers, leading to an increase in acidity and the loss of fertility (Esipisu 2011). If this is the case, we could be picking up on the fact that land more recently brought into a fertilizer rotation could experience higher gains from fertilizer use and that land with a long history of fertilizer use may no longer experience quite the same gains due to loss in inherent soil productivity. While this claim cannot be substantiated from my model, the pattern in response levels does mimic the process of fertilizer diffusion across the country.

Instead, I can compare the values I estimate with others found in the literature. Matsumoto and Yamano (2011) found marginal products varying across the western and higher potential regions between 11 and 20. Their analysis, however, excluded eastern Kenya where I find the highest returns. Marenya and Barrett (2009b) found the marginal product of nitrogen to be 17.64 for both Vihiga and South Nandi districts. While I estimate the value to be closer to 13.9, they did have a standard error of about 8, meaning my result is well within their confidence interval. Mbata (1997) looked at response to fertilizer in the Rift Valley, finding marginal products between 12 and 18, depending on the district. What the literature is lacking is fertilizer profitability work in eastern Kenya, meaning I am unable to compare my seemingly high estimates to any other previous work. In the absence of other estimates against which to compare, these values will be used in the remainder of the profitability analysis.

Chapter 5: Fertilizer Profitability Scenarios

Using the marginal and average products estimated in the last chapter, the profitability of fertilizer is assessed in this chapter under various relative price scenarios. The scenarios are chosen in order to compare how differences in the chosen relative price scenario change overall profitability metrics and to more accurately reflect the maize market status of households and the true opportunity costs of acquiring fertilizer and growing maize. Section 1 describes how the prices of nitrogen are calculated; section 2 describes how the prices of maize are determined; and section 3 includes the mechanics and results of the profitability scenario analysis at different levels of disaggregation.

1. Price of Nitrogen

Because applied fertilizer is split into its nutrient components in the production function, so too must be its price. I compute the price of nitrogen at the field level using the observed price paid by the household for DAP and CAN, the two most commonly used fertilizers in the dataset. I calculate the price per kilogram of nitrogen using the conversion factors found in Appendix 2 and the price paid by the farmer for that fertilizer type. Then, given variability in prices paid by households and in order to diminish any household level measurement error, I calculate the average market price of nitrogen at the district level. Similarly, by averaging at the district level, I automatically weight the prices of nitrogen by the relative frequency of DAP and CAN on fields; therefore, in places where more top dressing is used, the price of CAN will be weighted more heavily by design.

Again, these prices do not necessarily accurately reflect the cost of acquiring fertilizer, especially in places where fertilizer retail outlets may be few and far between or where infrastructure may be poorly developed. Given the stated importance of transactions and transport costs, I create an estimated transport cost, one essential component of the full gamut of transactions costs, from the household to the nearest fertilizer seller. In each survey year, I observe the distance (in kilometers) from the household to the nearest fertilizer seller, but only in 2010 do I know the cost of moving between the locations (via matatu, motorbike, bicycle etc.). I calculate a village-averaged transport cost per kilometer using the cost observed in 2010 and then multiply that value by a village-averaged distance to the nearest fertilizer seller in each year. This method assumes a linear relationship between the total transport cost and the distance traveled, with the transport cost per kilometer remaining constant over time. While transport costs per kilometer likely have varied between years, particularly in those areas where the distance to the nearest fertilizer seller has dropped dramatically, this method serves as a sensible way to estimate these unobservable costs.[13] These calculated transport costs are added to the district level market prices of nitrogen to arrive at what will hereafter be described as the "acquisition price" of fertilizer. Because the transport cost varies at the village level, this "acquisition price" too is specific to the village.

Table 16 shows the average distance from the household to the nearest fertilizer dealer. In general, distances are low and, in some zones (i.e., the three lowlands zones), falling

[13] For example, one thing the acquisition price does not account for is the fluctuations in fuel prices between years and the extent to which those changes contribute to changes in transportation cost. While I could adjust for changes in the relative price of fuel, a large number of households report using transportation that does not require fuel. For example, the largest percentage of households (about 15 percent) use bicycles to transport fertilizer. Assuming one mode of transport at the village level across would introduce a considerable amount of error. For these reasons, the transportation costs do not account for fluctuations in fuel prices between years. For reference, using yearly averaged fuel prices in Nairobi from the Central Bureau of Statistics in Kenya, I find the real price of fuel was 1.513 times the level in 2009/2010 for the 2007 survey (would have purchased fertilizer in 2006) and 1.552 times the level in 2009/2010 for the 2004 survey (would have purchased fertilizer in 2003).

considerably over time. As evidence from high standard deviations, however, slight increases and decreases should not be the focus, as variation within zones is immense. Instead, one should note the overall reduction in transportation necessary to access fertilizer over time. Both of these findings provide further justification for adding the cost of transportation to the price of fertilizer to arrive at a fertilizer acquisition cost specific to households at a given location and at a particular moment in time.

Table 16: Mean and standard deviation of distance from hh to nearest fertilizer seller (km)

	1997	2000	2004	2007	2010
Coastal Lowlands	25.2 (16.8)	23.8 (11.5)	17.3 (22.3)	8.7 (12.0)	4.6 (4.4)
Eastern Lowlands	10.1 (13.1)	6.0 (10.0)	3.7 (5.1)	2.8 (2.9)	3.6 (3.9)
Western Lowlands	16.2 (10.5)	12.5 (6.2)	7.0 (7.0)	3.9 (1.6)	4.3 (2.6)
Western Transitional	6.7 (5.9)	4.8 (5.4)	2.9 (2.4)	3.6 (3.1)	4.1 (2.8)
High Potential Maize	5.3 (8.2)	3.8 (3.9)	3.1 (3.2)	3.7 (3.6)	5.2 (4.3)
Western Highlands	3.3 (4.0)	1.8 (1.8)	1.3 (1.1)	2.3 (1.8)	2.9 (1.6)
Central Highlands	2.8 (3.9)	1.5 (1.6)	1.3 (0.8)	1.4 (1.4)	1.4 (1.5)
Marginal Rain Shadow	23.6 (8.3)	2.2 (1.9)	7.0 (9.7)	2.9 (2.7)	4.4 (5.1)
Total sample	7.5 (10.1)	5.9 (7.7)	3.8 (6.5)	3.4 (3.9)	4.0 (3.6)

Note: Households self-report distance to nearest fertilizer seller. Given the chance for measurement error, these values are averaged at the village level (in each survey year) for use in analysis.

Table 17 shows the average computed market and acquisition price of fertilizer in real 2010 terms. In some areas, the cost of transport creates a significant wedge between the market and acquisition prices of nitrogen (e.g., 1997 in the Coastal Lowlands). On average, though, the cost of transport adds between 50 and 100 KSH to the market price of fertilizer, particularly in the more recent survey years. Where the cost per kilogram of nitrogen is about 200 KSH, this represents a 25 to 50 percent increase in the cost of using fertilizer. The extent to which this additional fixed cost changes the profitability of purchasing and using fertilizer is assessed in this section.

Table 17: Mean price of nitrogen per kg (2010 prices, KSH)

	N price	1997	2000	2004	2007	2010
Coastal Lowlands	Market	407	230	216	227	258
	Acquisition	771	527	437	261	318
Eastern Lowlands	Market	344	246	217	189	166
	Acquisition	477	299	262	238	219
Western Lowlands	Market	632	450	376	315	234
	Acquisition	951	725	465	388	308
Western Transitional	Market	356	332	273	230	216
	Acquisition	456	378	303	263	258
High Potential Maize	Market	457	351	278	239	224
	Acquisition	507	392	307	273	266
Western Highlands	Market	519	367	247	254	205
	Acquisition	582	411	276	307	258
Central Highlands	Market	314	267	226	216	199
	Acquisition	378	308	267	254	243
Marginal Rain Shadow	Market	285	227	195	182	175
	Acquisition	600	272	236	211	215
Total sample	Market	432	337	268	242	213
	Acquisition	550	418	316	285	263

Note: Market prices reflect district averages. Acquisition prices reflect district level averaged market prices plus the village level calculated transport cost of fertilizer between households and the nearest fertilizer dealer. Prices are adjusted to 2010 levels using the CPI.

2. Price of Maize

While fertilizer prices and transport costs are known at the time of purchase and use, the price for which maize will sell on the market months later is not known to the farmer. Instead, the farmer makes expectations about what that price will be. Again, because I want to model what farmers *perceive* fertilizer profitability to be at the time of planting, it is necessary to predict and use what farmers might have expected maize prices to be at the time of harvest. I use expected maize selling prices using the results of a technique employed by Muyanga (forthcoming) of regressing the price at which farmers sell their maize at the end of the season on the information available to farmers at the time of planting and other factors that determine the price farmers receive. These include current and lagged NCPB (government maize board) prices, regional markets current and lagged prices, distances from the regional markets, and the type of buyer to which farmers normally sell their maize. With the regression estimates, I predict the selling price of maize farmers likely envisioned at the time of planting. With estimates at the household level, I average to the district level to remove any possible measurement and use these values as expected maize selling prices.

While the selling price of maize is the usual metric for calculating the marginal and average value product of output, a significant number of households in the dataset either do not sell their maize (i.e., keep for home consumption) or sell some maize but buy more making them overall net buyers. Table 18 shows the percent of net buyers and net sellers in this data set each year. The excluded group represents autarkic households or those that rely exclusively on gifts or aid. As with most other metrics, variation between zones can be substantial. Expectedly, households in the lower potential areas are more likely to be net buyers while households in higher potential areas are more likely to be net sellers. Differences between the survey years underscore the importance of variation in yields as a significant driver of maize market standing, which cannot necessarily be predicted at the start of the season.

For the majority of households, then, a better measure of the opportunity cost of growing maize might be its buying price. In the survey, households are asked how much they paid for maize grain, posho maize (i.e., unrefined maize flour) and packaged sifted maize meal, the three main ways of procuring maize to feed a household (apart from "green" maize, which is not considered here). Given the different levels of processing and packaging required for these three types of purchased maize, the prices can vary considerably between them. On average, there is between a 15 to 25 KSH per kilogram premium when purchasing already sifted maize flour. While recognizing the different levels of processing and prices at which maize can be procured in rural areas, I will use the price of maize grain as the buying price of maize in the profitability analysis. While not all households may have access to packages of shifted flour, maize grain should be available for purchase either from neighboring households or retail markets then processed either at home or a local mill.

Table 18: Percent net buyers and net sellers of maize by zone and year (autarkic excluded)

		1997	2000	2004	2007	2010
Coastal Lowlands	Net buyer	89	88	91	72	88
	Net seller	2	9	9	17	5
Eastern Lowlands	Net buyer	81	71	54	57	60
	Net seller	13	23	36	31	25
Western Lowlands	Net buyer	75	79	80	60	53
	Net seller	8	12	14	21	30
Western Transitional	Net buyer	77	57	41	32	37
	Net seller	13	34	44	50	42
High Potential Maize	Net buyer	25	26	20	19	36
	Net seller	62	60	71	73	46
Western Highlands	Net buyer	53	55	51	44	36
	Net seller	26	33	32	43	48
Central Highlands	Net buyer	63	52	52	40	53
	Net seller	21	39	33	46	23
Marginal Rain Shadow	Net buyer	80	88	52	26	44
	Net seller	7	13	45	41	34
Total sample	Net buyer	57	51	46	38	47
	Net seller	30	39	44	49	36

Note: Net buyer defined as a household which purchases more maize than they produce in a given year. Net sellers are defined as households which sell more maize than they purchase in a given year. Households with a balance of zero (autarkic) or ones in which rely exclusively on gifts or aid are the excluded percentage.

Again, I am interested in what farmers *perceive* the buying price to be at the end of the season, not what they actually encounter. Instead of modeling expected buying prices using the same regression method as expected selling prices, I calculate the difference between the expected and actual (observed) selling prices and add that difference to the district-averaged actual (observed) buying prices to arrive at a district-averaged expected buying price. Table 19 shows the calculated expected buying and selling price of maize. In general, the buying price of maize is between 5 and 10 KSH more than the selling price, with a much larger wedge in 2004 than the other two years. The difference between the selling and buying price of maize represents the premium for not producing enough maize to feed a household. If the gap between buying and selling prices of maize forces a switch in fertilizer use profitability, then the riskiness of maize production, in general, is likely a reason a household would choose not to fertilize.

Table 19: Mean expected selling and buying price of maize per kg (2010 prices, KSH)

		1997	2000	2004	2007	2010
Coastal Lowlands	Sell price	51.6	38.5	31.4	21.7	25.8
	Buy price	-	-	40.5	37.4	29.8
Eastern Lowlands	Sell price	37.1	33.9	27.6	20.0	18.9
	Buy price	-	-	34.9	25.8	27.1
Western Lowlands	Sell price	43.2	37.0	29.8	21.6	22.1
	Buy price	-	-	34.9	22.6	23.5
Western Transitional	Sell price	36.6	33.9	27.4	19.0	20.8
	Buy price	-	-	34.2	21.5	22.7
High Potential Maize	Sell price	37.7	33.0	26.9	18.0	20.5
	Buy price	-	-	33.9	20.4	23.0
Western Highlands	Sell price	40.1	37.6	30.4	21.9	22.0
	Buy price	-	-	35.3	23.2	22.3
Central Highlands	Sell price	42.6	37.0	28.9	21.0	20.4
	Buy price	-	-	34.7	24.1	26.3
Marginal Rain Shadow	Sell price	36.5	-	27.8	17.5	19.4
	Buy price	-	-	33.8	21.3	27.5
Total sample	Sell price	39.4	34.9	28.3	19.8	21.0
	Buy price	-	-	34.7	22.9	24.3

Note: Purchase prices of maize not observed in 1997 or 2000. Expected selling price modeled using a method by Muyanga (forthcoming). The difference between expected and actual selling prices are used to estimate expected buying prices. Prices are adjusted to 2010 levels using the CPI.

The one remaining value that I do not capture here is the distance a household needs to travel to either sell or purchase maize. While I do observe the distance a household traveled to make its largest maize sale in certain survey years, this variable does not necessarily capture the closest alternative for the household. A farmer could make the choice to travel a greater distance in order to make a larger sale, bypassing several other markets along the way, or simply to sell from the farm to other households in the village. In the 2010 dataset, over 50 percent of households claimed to sell their maize from the farm (the buyer came to them). Furthermore, I never observe how far a household needs to travel to purchase maize. For these reasons, the transport cost of selling and acquiring maize are not included here.

3. Fertilizer Profitability Scenarios

Using the various prices of fertilizer and maize described above, I estimate several profitability scenarios as summarized in Table 20. The first four scenarios represent different combinations of market and acquisition prices of nitrogen alongside selling and buying prices of maize. Then, in order to get a closer approximation of the true opportunity cost of maize production specific to the surveyed households, I estimate an additional profitability scenario (scenario five) where I choose the maize price based on observed net buying and selling behavior.

Table 20: Five fertilizer profitability scenarios

Profitability Scenario	Price of N	(Expected) Price of maize
1	Market price	Selling price
2	Market price	Buying price
3	Acquisition price	Selling price
4	Acquisition price	Buying price
5	Acquisition price	Buying or selling price, given maize market standing of household

Note: The "acquisition" price of fertilizer is the market price plus a calculated transport cost. Buying and selling prices of maize are calculated "expected" prices. All prices are averaged at the district level, expect for the acquisition price which includes a village-level transport cost.

If a household is a consistent net seller of maize across all five surveys (114 of 906 households), then the selling price of maize is attributed. If the household is a consistent net buyer of maize (131 of 906), then the buying price of maize is used. If the household is sometimes a net buyer and sometimes a net seller (661 of 906), then an average of the two is used. These values seek to mimic the household perception of the opportunity cost of producing maize by attributing the maize price that best matches their observed maize market standing over time.

Table 21: Relative price scenarios (nitrogen/maize per kilogram) over time by zone

	Relative price scenario	1997	2000	2004	2007	2010
Eastern Lowlands	1	9.5	7.3	7.9	9.3	8.7
	2	-	-	6.3	7.4	6.0
	3	13.3	8.8	9.5	11.7	11.4
	4	-	-	7.5	9.4	7.9
	5	-	-	8.2	10.1	8.9
Western Lowlands	1	14.5	12.2	12.8	14.6	10.6
	2	-	-	10.8	13.9	10.0
	3	21.7	19.3	15.8	18.0	13.9
	4	-	-	13.2	17.2	13.1
	5	-	-	14.2	17.5	13.4
Western Transitional	1	9.7	9.8	9.9	12.1	10.4
	2	-	-	7.9	10.7	9.6
	3	12.5	11.1	11.0	13.9	12.4
	4	-	-	8.8	12.3	11.4
	5	-	-	9.7	13.0	11.9
High Potential Maize	1	12.2	10.6	10.3	13.3	10.8
	2	-	-	8.1	11.8	9.7
	3	13.5	11.8	11.4	15.1	12.8
	4	-	-	9.0	13.4	11.6
	5	-	-	10.4	14.4	12.3
Western Highlands	1	12.9	9.8	8.0	11.6	9.3
	2	-	-	6.9	10.9	9.1
	3	14.4	11.0	8.9	13.9	11.7
	4	-	-	7.7	13.2	11.5
	5	-	-	8.2	13.5	11.5
Central Highlands	1	7.4	7.2	7.8	10.3	9.8
	2	-	-	6.6	8.9	7.6
	3	8.9	8.3	9.3	12.2	11.9
	4	-	-	7.7	10.5	9.2
	5	-	-	8.3	11.3	10.2
Total sample	1	11.1	9.9	9.7	12.4	10.1
	2	-	-	7.9	11.1	9.0
	3	13.8	12.0	11.1	14.5	12.5
	4	-	-	9.0	13.1	11.1
	5	-	-	10.0	13.8	11.7

Note: For information on how the five scenarios are defined, see Table 20. Buying price of maize not observed in 1997 and 2000, so scenarios 2, 4, and 5 not estimated for these years.

Before looking specifically at the profitability calculations, it is useful to conceptualize relative prices under each of the aforementioned scenarios. As mentioned previously, what is more important than changes in the price of fertilizer and maize is the change in relative prices over time which better describes the incentive to use fertilizer (without, yet, including how applying fertilizer contributes to increased maize yields). Table 21 shows the relative price of fertilizer to maize (i.e., the inverse of what is used in the MVCR and AVCR calculation) under the five different profitability scenarios previously mentioned. As expected, these ratios vary tremendously between scenario, year and zone. A lower ratio signals that the incentive to use fertilizer is greater: the cost of the input is relatively less than the price of the output.

Over the total sample, I find that adding the transport cost of fertilizer increases the relative price of nitrogen to maize by about 25 percent (i.e., scenario three versus one; scenario four versus two). Using the selling price of maize as opposed to the buying price of maize makes nitrogen more expensive by about 15 percent (i.e., scenario two versus one; scenario three versus four) given that the buying price of maize is generally higher than its selling price. Expectedly, scenario five, where household maize market interactions are considered, is essentially an average of scenarios three and four given the distribution of net buyer and seller behavior. Temporally, trends are telling as well. Overall, these ratios do not show an overwhelming decline in the relative price of fertilizer to maize over time. Market prices of nitrogen were relatively high in 1997 but declined in 2000 and 2004. In 2007, the price of nitrogen increased much more than the cost of maize, forcing the relative price back up again. By 2010 the relative price had fallen again, but still not in line with 2000 and 2004 levels. This trend is somewhat amplified when adding the transport cost of fertilizer; the decrease in distance traveled to fertilizer retailers over time has steadily decreased the acquisition price of nitrogen. Notice how very high the relative prices were in 1997 under scenario three in the Western Lowlands. In general, the buying and selling prices of maize have moved together, with the buying price generally above the selling price no matter the year.

When focusing specifically on the 2010, zonal differences are still apparent. The lowest nitrogen to maize ratios are found in the Eastern Lowlands and Western and Central Highlands. In the lowlands, this is explained by relatively higher prices of maize, both buying and selling. In the highlands, nitrogen prices are, in fact, lower than in other parts of the country. The highest relative prices are found in the Western Lowlands, were nitrogen prices are highest and fertilizer use at its lowest. In the highest potential areas, the ratio hovers around 12, consistent with other work in the area. For example, Matsumoto and Yamano (2011) use a value of 13 during the years in their sample across western and central Kenya.

From there, the marginal value cost ratios (MVCRs) and average value cost ratios (AVCRs) are calculated under the same five relative price scenarios (refer back to equations 9 and 10 on page 17). All calculations are made using the marginal and average product at the district and soil group level by year (see Table 15) with prices of maize and fertilizer at the district level and transport costs at the village level (see Table A.7 of Appendix 7). The averages of these values by zone, year and scenario are shown in Table 22. Given there are several moving components (i.e., price of fertilizer and maize, transport cost, marginal and average products), the actual MVCRs and AVCRs are the best metric for summarizing overall expected profitability of fertilizer. Because of this, variation over time, space and profitability scenario produce interesting results. Recall that AVCRs measure the total gain in household income from using a unit of fertilizer (i.e., the average gain per unit) while MVCRs measure the gain in income from the last unit of fertilizer. Moreover, AVCRs give a sense of overall profitability while MVCRs relate to the profitability of a given level and can make statements about room for profitable expansion. A value of greater than one means the input choice is profitable, while a value of greater than two is considered profitable enough for risk averse farmers to want to use.

Table 22: Mean MVCRs and AVCRs for nitrogen to maize by profitability scenarios

Zone	Profit. scenario	MVCR					AVCR				
		1997	2000	2004	2007	2010	1997	2000	2004	2007	2010
Eastern Lowlands	1	4.55	5.79	4.89	4.62	4.36	4.62	6.03	5.80	5.42	5.69
	2	-	-	6.10	5.73	6.31	-	-	7.22	6.73	8.17
	3	3.62	4.81	4.03	3.64	3.26	3.68	5.01	4.79	4.28	4.24
	4	-	-	5.03	4.56	4.74	-	-	5.97	5.35	6.13
	5	-	-	4.71	4.23	4.21	-	-	5.59	4.97	5.46
Western Lowlands	1	1.76	2.50	2.09	1.97	2.39	1.95	3.82	2.54	2.36	3.30
	2	-	-	2.82	2.09	2.62	-	-	3.42	2.51	3.61
	3	1.42	1.74	1.67	1.75	1.79	1.43	2.10	2.02	2.09	2.46
	4	-	-	2.25	1.85	1.96	-	-	2.73	2.22	2.70
	5	-	-	1.98	1.81	1.89	-	-	2.40	2.16	2.60
Western Transitional	1	1.47	1.48	1.36	1.01	1.17	1.68	1.71	1.67	1.30	1.51
	2	-	-	1.70	1.14	1.27	-	-	2.09	1.47	1.64
	3	1.19	1.31	1.24	0.89	0.98	1.38	1.52	1.52	1.14	1.27
	4	-	-	1.55	1.01	1.07	-	-	1.91	1.29	1.39
	5	-	-	1.40	0.95	1.03	-	-	1.72	1.22	1.33
High Potential Maize	1	1.15	1.20	1.13	0.86	1.02	1.38	1.52	1.48	1.18	1.41
	2	-	-	1.42	0.99	1.15	-	-	1.86	1.35	1.57
	3	1.04	1.07	1.02	0.75	0.87	1.25	1.36	1.33	1.02	1.19
	4	-	-	1.28	0.86	0.97	-	-	1.68	1.17	1.33
	5	-	-	1.12	0.79	0.91	-	-	1.46	1.08	1.25
Western Highlands	1	1.22	1.55	1.77	1.29	1.62	1.34	1.69	2.07	1.50	1.99
	2	-	-	2.05	1.39	1.62	-	-	2.39	1.61	1.99
	3	1.09	1.39	1.60	1.11	1.30	1.20	1.52	1.86	1.28	1.60
	4	-	-	1.84	1.19	1.30	-	-	2.15	1.38	1.60
	5	-	-	1.73	1.15	1.30	-	-	2.01	1.33	1.60
Central Highlands	1	2.83	2.70	2.41	2.01	2.07	3.14	2.98	2.70	2.21	2.31
	2	-	-	2.89	2.31	2.67	-	-	3.24	2.54	2.99
	3	2.35	2.36	2.04	1.70	1.69	2.61	2.60	2.29	1.88	1.89
	4	-	-	2.44	1.95	2.19	-	-	2.74	2.15	2.26
	5	-	-	2.25	1.83	1.98	-	-	2.52	2.01	2.22
Total sample	1	1.82	2.13	1.94	1.49	1.70	2.03	2.39	2.34	1.83	2.19
	2	-	-	2.39	1.73	2.05	-	-	2.89	2.11	2.63
	3	1.53	1.82	1.67	1.26	1.36	1.71	2.06	2.02	1.54	1.75
	4	-	-	2.07	1.46	1.64	-	-	2.50	1.78	2.10
	5	-	-	1.88	1.37	1.52	-	-	2.28	1.67	1.95

Note: See Table 20 for information on how the five scenarios are defined. See equations 9 and 10 for the MVCR and AVCR formulas. See text for additional information on calculations.

Over the total sample, MVCRs are between 1.25 and 2.39 and AVCRs are between 1.50 and 2.89, depending on the year and relative price scenario. All of these values are over one, meaning fertilizer use is profitable across the sample, and sometimes over two, meaning quite profitable. The relatively small range in MVCRs and AVCRs means that even when considering the differences in buying and selling prices and maize and even when accounting for the transport cost of fertilizer, the profitability of applying nitrogen to maize does not vary wildly. Nowhere does adding those real costs suddenly make fertilizer use unprofitable. Instead, in the years where fertilizer use is unprofitable, it remains so across the various scenarios.

The highest MVCRs and AVCRs are found in the Eastern Lowlands due to high marginal and average products. With values between four and six, this suggests vast increases in household income from the use of fertilizer per unit and that the last unit of fertilizer was very profitable, implying that households could profitably use more. Fertilizer use is next most profitable in the Central Highlands where, again, both the average and last unit of fertilizer were particularly profitable (in most cases, with AVCRs and MVCRs over two). Interestingly, the zone where fertilizer is least profitable, on average, is the High Potential Maize zone where AVCR values are above one but MVCR values are either at one, slightly above or slightly below. This indicates that while profitable to use, households are likely using at or near the most profitable rates and that there would not be substantial gains from increasing dosage. In fact, in some cases, a decrease in the amount of fertilizer applied might be the most profitable strategy.

Values in Table 22 represent averages across an entire agro-ecological zone. Given rain stress levels, soil type, and past fertilizer application rates can vary within, I also calculate profitability levels by district and soil group, the level of disaggregation used in the remainder of the profitability analysis. These values can be found in Table 23 for profitability scenario five, which most closely approximates village and household level variation in relative prices. This level of disaggregation shows that within-zone variation is important. For example, Narok and Bomet districts, both found in the Rift Valley and High Potential Maize Zone, have substantially different profitability measures. The values found here will be used in the subsequent chapters where I analyze fertilizer use patterns alongside profitability levels.

Table 23: MVCRs and AVCRs (scenario five) by district, soil group and year

Province	District	Soil group	MVCR 2004	MVCR 2007	MVCR 2010	AVCR 2004	AVCR 2007	AVCR 2010
Eastern	Machakos	3	3.60	3.69	3.27	4.18	4.21	4.34
	Makueni	3	5.09	4.01	3.29	5.88	4.69	4.83
	Meru	1	2.21	1.51	1.46	2.48	1.70	1.67
	Mwingi	2	5.00	5.05	6.13	6.91	6.07	6.36
	Mwingi	3	4.22	4.38	5.91	4.42	5.13	8.17
Nyanza	Kisii	2	2.02	1.49	1.72	2.42	1.69	2.06
	Kisii	4	2.05	1.40	1.48	2.32	1.61	1.82
	Siaya	3	1.97	1.78	1.96	2.34	2.07	2.63
	Siaya	4	2.00	1.84	1.69	2.53	2.31	2.51
Western	Bungoma	2	2.11	1.27	1.54	2.48	1.64	1.90
	Bungoma	3	0.80	0.63	0.71	1.25	0.89	1.03
	Bungoma	4	1.57	1.02	0.96	2.06	1.38	1.38
	Kakamega	2	1.11	1.17	1.04	1.72	1.56	1.58
	Kakamega	3	0.97	0.66	0.76	1.30	0.96	1.11
	Kakamega	4	1.49	1.17	1.30	1.64	1.32	1.44
	Vihiga	3	0.93	0.55	0.64	1.15	0.69	0.83
	Vihiga	4	1.47	0.79	0.98	1.72	0.91	1.20
Central	Muranga	1	2.32	2.04	2.20	2.54	2.16	2.52
	Muranga	4	2.51	2.12	2.44	2.63	2.26	2.55
	Nyeri	1	2.06	1.92	1.90	2.44	2.13	2.16
	Nyeri	2	2.71	2.35	2.52	2.90	2.50	2.69
Rift Valley	Bomet	1	1.26	0.88	0.97	1.42	0.98	1.11
	Nakuru	1	0.71	0.54	0.61	0.92	0.68	0.85
	Nakuru	2	1.10	0.59	0.81	1.30	0.72	1.00
	Nakuru	3	1.05	0.69	0.84	1.24	0.79	1.03
	Narok	1	0.48	0.23	0.31	0.65	0.30	0.44
	Trans Nz.	4	1.04	0.98	1.18	1.51	1.48	1.68
	Uasin Gis.	1	0.99	0.61	0.69	1.34	0.87	0.98
	Uasin Gis.	2	1.34	0.82	0.88	1.84	1.30	1.32

Note: See Table 20 for information on how the five scenarios are defined. See equations 9 and 10 for the MVCR and AVCR formulas. See text for additional information on calculations. For reference, averages of variables included in these calculations can be found in Table A.7 of Appendix 7. The values found in this table are used in the profitability analysis that follows.

The MVCR and AVCR values described above represent measures of relative profitability, relying on the ratio of nitrogen to maize prices under different price specifications. Given the relative price of nitrogen to maize has not changed tremendously over the survey years (see Table 21), the relative profitability of fertilizer, as embodied in the MVCRs and AVCRs, show that changes over time are not substantial. Because the absolute levels of fertilizer and maize prices have, in fact, changed over time, an absolute level of profitability provides insight into the actual returns to fertilizer experienced in a given year. For example, Figure A.3 shows how the absolute prices of nitrogen and maize have moved with respect to the relative price of nitrogen and maize over time and by zone. Indexed to 1997 levels, these plots provide a sense of how the relative price stayed fairly constant despite decreases in prices over time.

Table 24 shows the net gain to the last kilogram of nitrogen applied (see equation 15) using the acquisition price of nitrogen and both the selling and buying price of maize for comparison. When using the selling price of maize, the net gain to the last unit of fertilizer application has diminished considerably across time. Even in Eastern Province, where the reduction in transport cost over time has reduced the most, the net gain to fertilizer use over time has decreased. The negative values in some of the more productive regions are a function of both lower marginal products of nitrogen (around or below 1, in most cases) and the prices of nitrogen and fertilizer. While nitrogen prices and transport distances are lowest in these areas, the selling price of maize, too, is relatively low, making the net gain to the last unit of nitrogen applied not particularly profitable.

Using the buying price of maize produces higher net gains given the buying price of maize is generally higher. This does not mean, however, that the net gains to net buyers of maize are greater than those to net sellers given net buying households need to purchase both maize and nitrogen (both cash outflows) while net sellers sell maize but purchase nitrogen (cash inflow and outflow). What this does show, though, is that (1) the net gain to fertilizer use is higher when using the price at which most household purchase maize and (2) the decrease in net gain to the last unit of fertilizer between 2004 and 2010 has been more severe when using the net buying price as opposed to the net selling price.

Table 24: Net gain to last kilogram of fertilizer applied (KSH) by district, soil group, year

Province	District	Soil group	Selling price of maize					Buying price of maize				
			1997	2000	2004	2007	2010	1997	2000	2004	2007	2010
Eastern	Machakos	3	1141	1111	603	504	433	-	-	837	904	938
	Makueni	3	940	1019	756	572	341	-	-	989	643	511
	Meru	1	311	327	249	96	68	-	-	370	177	178
	Mwingi	2	1350	1418	1031	815	719	-	-	1379	1276	1202
	Mwingi	3	1103	1185	818	666	683	-	-	1111	1063	1147
Nyanza	Kisii	2	77	272	231	123	164	-	-	305	160	159
	Kisii	4	64	207	226	90	115	-	-	296	122	110
	Siaya	3	379	454	338	291	265	-	-	637	333	321
	Siaya	4	414	386	391	340	206	-	-	721	387	259
Western	Bungoma	2	312	254	239	44	97	-	-	386	94	168
	Bungoma	3	-92	-38	-67	-106	-89	-	-	-4	-77	-50
	Bungoma	4	155	183	109	-16	-34	-	-	221	24	15
	Kakamega	2	182	93	15	34	7	-	-	91	67	18
	Kakamega	3	42	5	-43	-106	-67	-	-	21	-86	-58
	Kakamega	4	5	170	94	24	67	-	-	188	55	81
	Vihiga	3	-21	-36	-55	-168	-123	-	-	-9	-164	-111
	Vihiga	4	143	95	96	-84	-17	-	-	168	-78	1
Central	Muranga	1	507	424	304	221	169	-	-	401	265	332
	Muranga	4	704	560	378	265	226	-	-	490	317	423
	Nyeri	1	583	411	245	200	161	-	-	331	253	245
	Nyeri	2	866	690	442	314	293	-	-	567	381	404
Rift Valley	Bomet	1	154	72	89	-35	-23	-	-	121	-57	-2
	Nakuru	1	-146	-85	-113	-146	-130	-	-	-54	-114	-97
	Nakuru	2	-98	48	-11	-147	-78	-	-	85	-107	-34
	Nakuru	3	-56	26	-22	-111	-67	-	-	69	-68	-22
	Narok	1	-330	-143	-192	-243	-205	-	-	-126	-251	-187
	Trans Nz.	4	95	-58	-17	-17	15	-	-	59	17	71
	Uasin Gis.	1	-80	15	-19	-101	-80	-	-	41	-78	-77
	Uasin Gis.	2	-18	50	50	-56	-32	-	-	133	-27	-28
Total sample			220	269	186	61	78	-	-	303	113	144

Note: See equation 15 for calculation. The acquisition price of nitrogen is used throughout. Buying price of maize not observed in 1997 and 2000.

To summarize, this analysis shows that using different relative price scenarios in fertilizer profitability analysis does produce small differences in the level of profitability but never changes the overall level from profitable to unprofitable. For example, adding the transport cost of fertilizer in recent years never forces an otherwise profitable input use decision based on market prices alone to become unprofitable. Furthermore, while there are generally differences in expected buying and selling prices of maize, the inclusion of either in the profitability analysis never produces overwhelmingly different results. This suggests that farmers with different interactions with the maize market (net buyer versus net seller) do not encounter wildly disparate profitability measures.

Relative profitability measures across years are driven by changes in relative prices and expected responsiveness to inorganic fertilizer. As expected, 2007 was the least profitable year for using fertilizer while 2000 was the most profitable. Differences in values across space, again, are the result of different marginal and average products of nitrogen and differences in relative prices. I find much more substantial variation across space than I do over time. Moreover, absolute profitability measures show a decline in the net gain from the last unit of fertilizer over time. While relative nitrogen to maize prices have not changed considerably over the survey years here, the absolute prices of fertilizer and maize have moved such that the absolute profitability of fertilizer has declined. In some areas (Eastern Lowlands), expanding fertilizer use appears to be a profitable strategy while in others (High Potential Maize Zone) fertilizer use appears at or even slightly beyond optimal levels. I investigate these findings alongside actual use patterns in the next chapter.

Chapter 6: Fertilizer Profitability and Use Decisions

In this chapter, I investigate fertilizer use decisions alongside the profitability metrics estimated in the last chapter and calculated optimal application rates in an attempt to uncover where fertilizer use could be profitably expanded and to understand the differences between households using and not using fertilizer at calculated profitability levels. Section 1 includes summary statistics that compare profitability metrics to actual fertilizer use levels; section 2 describes estimated optimal fertilizer application rates under two different risk scenarios; section 3 considers the size of the "gap" between calculated optimal fertilizer application rates and observed use levels; and section 4 shows the additional revenue from fertilizer use at current and optimal levels.

1. Summary Statistics on Fertilizer Profitability and Use

In this section, I compare the profitability work from the last chapter to observed fertilizer use rates. While various scenarios were run, the results from scenario five are used in the remainder of the analysis given they capture the dynamics most closely attributable to household specifics. Table 25 shows the MVCR and AVCR from profitability scenario five by district, soil group and year alongside the percent of maize fields where fertilizer was used (in any amount) and the average application rate of fertilizer user. I only include the three most recent survey years given data availability and the desire to focus on fertilizer decisions in the recent past. For more on use rates in all five survey years, see Table A.9 in Appendix 8.

From this table, we learn a lot about where there appears to be room for fertilizer use expansion, both in the percent of fertilized fields and application rates, and where households are likely using fertilizer at or above the most profitable levels. Spatially, there is incredible variation across the country and, even over a short seven year period, instances of considerable change in the number of households using fertilizer. Recall from the production function discussion that standard errors were relatively high in estimation, meaning profitability values should not be interpreted as precise, but instead used as a guide to understanding an overall picture.

Table 25: MVCR, AVCR, and actual fertilizer use rates by district, soil group, and year

Province	District	Soil group	2004				2007				2010			
			MVCR	AVCR	% use fert	N per ha	MVCR	AVCR	% use fert	N per ha	MVCR	AVCR	% use fert	N per ha
Eastern	Machakos	3	3.60	4.18	58	13	3.69	4.21	67	11	3.27	4.34	80	21
	Makueni	3	5.09	5.88	77	10	4.01	4.69	70	16	3.29	4.83	81	25
	Meru	1	2.21	2.48	95	25	1.51	1.70	90	28	1.46	1.67	89	30
	Mwingi	2	5.00	6.91	4	22	5.05	6.07	11	13	6.13	6.36	19	29
	Mwingi	3	4.22	4.42	29	3	4.38	5.13	14	13	5.91	8.17	30	23
Nyanza	Kisii	2	2.02	2.42	100	37	1.49	1.69	100	28	1.72	2.06	97	39
	Kisii	4	2.05	2.32	99	23	1.40	1.61	100	26	1.48	1.82	97	41
	Siaya	3	1.97	2.34	9	9	1.78	2.07	28	7	1.96	2.63	33	20
	Siaya	4	2.00	2.53	20	11	1.84	2.31	47	12	1.69	2.51	38	36
Western	Bungoma	2	2.11	2.48	96	34	1.27	1.64	95	51	1.54	1.90	93	42
	Bungoma	3	0.80	1.25	79	57	0.63	0.89	100	41	0.71	1.03	100	43
	Bungoma	4	1.57	2.06	96	48	1.02	1.38	93	54	0.96	1.38	93	56
	Kakamega	2	1.11	1.72	97	72	1.17	1.56	93	55	1.04	1.58	100	67
	Kakamega	3	0.97	1.30	67	49	0.66	0.96	78	52	0.76	1.11	81	51
	Kakamega	4	1.49	1.64	58	26	1.17	1.32	75	25	1.30	1.44	63	22
	Vihiga	3	0.93	1.15	71	28	0.55	0.69	87	28	0.64	0.83	86	34
	Vihiga	4	1.47	1.72	100	25	0.79	0.91	93	24	0.98	1.20	94	35
Central	Muranga	1	2.32	2.54	89	22	2.04	2.16	93	15	2.20	2.52	81	38
	Muranga	4	2.51	2.63	100	12	2.12	2.26	75	18	2.44	2.55	50	17
	Nyeri	1	2.06	2.44	97	37	1.92	2.13	96	26	1.90	2.16	96	30
	Nyeri	2	2.71	2.90	73	25	2.35	2.50	63	27	2.52	2.69	53	35
Rift Valley	Bomet	1	1.26	1.42	100	21	0.88	0.98	100	19	0.97	1.11	100	22
	Nakuru	1	0.71	0.92	95	24	0.54	0.68	94	23	0.61	0.85	85	35
	Nakuru	2	1.10	1.30	81	23	0.59	0.72	67	23	0.81	1.00	50	19
	Nakuru	3	1.05	1.24	98	22	0.69	0.79	98	17	0.84	1.03	96	25
	Narok	1	0.48	0.65	24	13	0.23	0.30	53	9	0.31	0.44	18	16
	Trans Nz.	4	1.04	1.51	92	55	0.98	1.48	90	60	1.18	1.68	94	53
	Uasin Gis.	1	0.99	1.34	92	36	0.61	0.87	91	43	0.69	0.98	94	40
	Uasin Gis.	2	1.34	1.84	95	51	0.82	1.30	96	64	0.88	1.32	98	56

Note: MVCR and AVCR levels based on profitability scenario five (see Table 15). The "percent use fert" column shows the percent of fields where fertilizer was applied at any level. The "N per ha" column shows the average nitrogen application rate by fertilizer users. For more on use rates across all survey years, see Table A.9 in Appendix 8.

As mentioned previously, the areas with the highest MVCRs and AVCRs are in the Eastern Lowlands, comprising most districts (Machakos, Makueni, Mwingi) in the Eastern Province, and Western Lowlands, meaning Siaya district in Nyanza Province. These areas also happen to have the lowest percentage of fertilizer users and the lowest dosage rates, particularly in earlier years. What I do find in these districts is an increase in the percentage of fertilized fields and an increase in the amount of nitrogen per hectare applied by fertilizer users over these three survey years. This suggests that the gap between where it is profitable to use and what households are actually doing has narrowed between 1997 and 2010, although more so in the Eastern Province than the lowlands areas of Nyanza (Siaya district). There is likely still room for expansion of fertilizer use in these lowlands areas of Kenya but, in the absence of other research against which to corroborate, further household level research should be conducted before prescribing fertilizer use at higher levels. Recall, also, that the lowlands areas have the lowest rainfall levels and highest rain stress, making maize production and fertilizer use particularly risky. Because of this, households might require a higher MVCR and AVCR before deciding to use fertilizer in order to account for the risk involved.

The next highest MVCR and AVCR levels are found in the Central Highlands (Central Province and Meru district in Eastern Province) where fertilizer use levels are considerably higher than the last group. Recall, however, the lack of significance and concavity in the squared term for this zone group, meaning these values should be interpreted with caution. Within this category, there appears to be a divide between areas with volcanic soils (soil group one) and other soil types. Those with volcanic soils are more likely to use fertilizer (around 90 percent) and at higher levels. The MVCRs on the non-volcanic soils are higher, though, suggesting that fertilizer use could be profitably expanded in these areas. Kisii district in the Western Highlands has some of the most constantly fertilized fields and, furthermore, at the highest levels. MVCR levels here suggest households are likely applying somewhere around the appropriate levels. There has not been a noticeable increase in the amount of fertilizer used in either of these highlands areas over the survey years in question, as fertilizer has been a more constantly applied input over a much longer period of time.

The remaining zones are the High Potential Maize and Western Transitional Zones, comprising all of Western and Rift Valley Provinces. Here, I find the lowest MVCRs and AVCRs across the board. On average, households see a gain in household income from using fertilizer (AVCR>1), however the last unit is generally at break-even profitable levels (MVCR=1) or not profitable at all (MVCR<1), meaning those households using fertilizer could be doing so at optimal or slightly more than optimal levels. There are some areas of Nakuru and Narok districts (Rift Valley) where fertilizer use does not appear profitable (AVCR<1). We do find relatively lower levels of fertilizer use in some of these areas (Narok), although some households appear to make making the non-profitable choice to use fertilizer on maize fields. Nakuru may be a case where we are not picking up on some important agro-ecological characteristic that makes farmers want to use fertilizer; while we find it unprofitable to apply, households are still applying at very high levels. Overall, households in these higher potential areas seem to have approached levels of optimality in fertilizer use and, as suggested in some of these values and in other research about the occurrence of high soil acid levels, perhaps more than optimal in some cases.

2. Optimal Nitrogen Use Rates

To build on the conclusions from the last section, I estimate disaggregated optimal nitrogen application rates using the production function estimates to compare with observed use patterns. Useful fertilizer application recommendations should be grounded in observed response rates and consider the local environment in which farmers operate. As such, I estimate economically optimal fertilizer application rates at the district and soil group level and use observations from all five years of survey data to hone in on more accurate levels. Optimal nitrogen application rates are estimated under two different scenarios: (1) where the MVCR=2 and (2) where the MVCR=1. Technically speaking, the economic optimal level of nitrogen for a risk neutral household would be where the MVCR=1, however, I also am interested in how a risk averse household should operate and, therefore, use a value of two, where a risk premium is added, as well. By rearranging these equations, the optimal nitrogen application rate is found by equating the marginal physical product (MPP) of nitrogen with two times the nitrogen to maize price ratio for the risk averse household and one times the nitrogen to maize price ratio for the risk neutral household. The marginal physical product is obtained by taking the first derivative of the production function, equation 19, with respect to nitrogen, which yields:

$$MPP_N = \partial y/\partial N = \beta_1 zone + 2\beta_2 N{*}zone + \beta_3 P{*}zone + \beta_4 soil + \beta_5 rain{*}zone \qquad (20)$$

When setting equation 24 equal to the price of nitrogen over the price of maize and solving for N, this turns into:

$$N^* = \frac{(1+\rho)(P_N/P_M) - \beta_1 zone - \beta_3 P{*}zone - \beta_4 soil - \beta_5 rain{*}zone}{2\beta_2{*}zone} \qquad (21)$$

where the 1+ρ term in front of the price ratio embodies the risk assumption, meaning equals one for the risk neutral case and two for the risk averse case. Using the production function coefficients (represented by β_i); the price ratios used in profitability scenario five (i.e., acquisition price of nitrogen and expected maize price specific to household maize market standing); expected rain stress conditions; and observed levels of phosphorous use, I calculate the optimal nitrogen application rate at the field level then average across all observations in each district and soil group to arrive at district and soil group level optimal nitrogen application rates.

It is important to note that calculations of optimal input use levels are derived under the assumption that all other inputs remain at their same levels. For example, due to the specification of the production function, optimal nitrogen application rates are a function of various other inputs, including phosphorous. This means that the amount of fertilizer applied (or the portion of it that was phosphorous) by the household during one of the survey years will influence the calculated optimal amount of nitrogen, making past fertilizer application endogenous to the optimal fertilizer application problem.

Calculated optimal nitrogen application rates are found in Table A.8 of Appendix 9 under both the MVCR=1 and MVCR=2 scenarios. The first thing to note is the size of the standard deviations. Across the sample, irrespective of district, standard deviations are very large, sometimes larger than the average. The overall coefficient of variation on optimal fertilizer use levels is about 70 percent over the total sample, with tremendous variation across space. This should not come as much of a surprise given high standard deviations in actual observed application rates (also in Table A.8) and high standard errors in the production function. In general, there is a lot of variation in the system, not all of which can be explained by the observable variables. With this observation, estimated optimal levels should be interpreted with caution; there are probably a number of household or field level considerations to make before recommending a particular level of fertilizer use.

Secondly, there are several instances in this table where the optimal nitrogen use rate is zero or very close to zero, meaning there is no positive rate of nitrogen use that satisfies the MVCR=1 or MVCR=2 requirement.[14] This happens when the marginal product of nitrogen is relatively low and/or the relative nitrogen to maize price is relatively high. Given these values, there is no level of nitrogen that would produce MVCR levels in line with the chosen scenarios. Note that on the volcanic soils in Narok, this happens under both the MVCR=1 and MVCR=2 scenarios, congruent with earlier findings that fertilizer use is not profitable in this area. Furthermore, a number of districts in both the Rift Valley and Western Province have zero or near zero values under the MVCR=2 scenario, meaning there are many field level observations where there is no positive value of nitrogen application that produces an MVCR=2, although there are positive values that satisfy the MVCR=1 requirement.

Finally, differences across space, as usual, are telling. In the lowlands areas, there is a difference of less than ten between the MVCR=1 and MVCR=2 scenarios. In the Eastern Province, optimal values are between 20 and 30 while in Siaya, values are between 10 and 25. In the highlands and high potential maize areas, the difference between the two scenarios is much greater. This is the product of less concavity in the production function for these areas, meaning the estimated optimal levels are very sensitive to the chosen relative price scenario. Furthermore, rates under the MVCR=1 scenario are quite high in the highlands areas, particularly in the Central Province. In the Rift Valley and Western Province, estimates seem more reasonable under the MVCR=1 scenario.

[14] Negative estimated optimal fertilizer application rates at the field level are replaced with zeros before averaging to the district and soil group level.

3. The "Gap" Between Optimal and Observed Fertilizer Use Levels

In this section, I investigate the size of the "gap" between calculated optimal application levels and what is currently being used by farmers. In Table A.8 and Table A.9 of Appendix 8, I show the actual observed levels of nitrogen use by survey year alongside the estimated optimal use levels described in the last section. Despite the very high standard deviations, in both estimated and observed nitrogen application values, this model produces optimal nitrogen application levels often very similar to what we observe households using, another check to its credibility. In some cases, estimates seem unreasonably high (e.g., much of Central Province) or low (e.g., the third soil group in Vihiga district), but this is not the norm. High standard deviations have been a feature of this analysis throughout; because of this, optimal values or the gap between them and actual observed should not be interpreted as absolute.

Furthermore, it is quite clear that use levels have changed dramatically in some areas over the survey years. 2010 levels are much more in line with estimated optimal levels, another indication that learning is taking place. This trend is particularly true in Eastern Province and Siaya. For example, in Machakos, fertilizer use levels were 4 kg/ha in 1997 as compared with 21 in 2010. In parts of Siaya, fertilizer application rates went from 0 in 1997 to about 20 kg/ha in 2010. Given estimated optimal values between 30 and 35 kg/ha, farmers are fast approaching calculated optimal levels. Moreover, with the high standard deviations observed, some fertilizer users might actually be using near optimal levels already. If these numbers are to be interpreted as exact, fertilizer users should increase their fertilizer application rates by about 10 kg/ha. These areas, however, are the ones with highest rain stress (lowest rainfall) and the greatest risk to maize production. Before suggesting farmers increase their investments in fertilizer use, one should consider risk preferences in addition to profitability measures.

The high potential and transitional areas in the Rift Valley and Western Provinces have a slightly different picture. Table A.8 shows that households are, on average, at or above the estimated profitable rates (where the MVCR=1). These findings are congruent with findings from earlier in this chapter where I note that the marginal value cost ratio is at or below one and in line with findings by (Matsumoto and Yamano 2011). In these areas, there is, on average, no gap to be filled. Instead, the average farmer appears to be using above optimal levels and could increase income by using less. On the volcanic soils in Uasin Gishu, for example, households in 2010 were using almost 20 kg/ha too much on their fields. Interestingly, average application rates in 1997 were much closer to estimated optimal levels. These areas also exhibit somewhat different fertilizer trends over time. In some districts (Uasin Gishu, Trans Nzoia, and some areas of Kakamega and Bungoma), fertilizer use has increased by 10-20 kg/ha between 1997 and 2010 although not always towards the most profitable levels; in some of these same areas, 1997 levels were more profitable than 2010 ones. In other districts, (Bomet, Nakuru, and the other parts of Kakamega and Bungoma), average fertilizer use values have remained fairly steady, with some areas still far above what I estimate are profitable levels of fertilizer use (Nakuru). In general, these districts of Kenya represent areas where expanding fertilizer use would not be a profitable strategy. In fact, further analysis should be conducted into the likely overuse (a gap in the other direction) of inorganic fertilizer.

The last area of focus is the highlands. Given generally high optimal use levels, it is difficult to compare the size of the gap between these values and actual ones with confidence. These unreasonably high estimates (a function of the low level of concavity in the production function described earlier) are clearly the reason for the MVCR and AVCR levels described earlier in this chapter. With estimated optimal fertilizer application rates around 70 kg/ha, it is no wonder that MVCR levels would be around two for households already applying around 30 kg/ha on average. Seeing these (likely) unreasonably high estimated optimal levels in these areas, a result of an unimpressive level of concavity in the production function, provides further justification to the fact that households in the highlands areas might already be using close to optimal levels. The one area that estimated optimal levels are near actual observed levels is in Vihiga district in Western Province.

In summary, analyzing the "gap" between optimal fertilizer application rates and observed use levels provides further evidence to the claim that households in lowlands areas are quickly approaching optimal levels of fertilizer use and that households in the high potential and transitional areas are likely at or beyond the most profitable levels. Furthermore, unreasonably high estimated levels in the highlands call into question the accuracy of the MVCR levels estimated for these areas, leading one to ambiguously believe that households could be applying somewhere near optimal levels already.

4. Revenue Added from Fertilizer Use at Current and Optimal Levels

In this section, I return to the discussion of absolute levels of fertilizer profitability. Here, I calculate the revenue added through the use of nitrogen, both at observed use rates and at estimated optimal levels. This calculation is not an average or at the margin, but instead a measure of the value of the additional output provided by fertilizer use minus the cost of fertilizer at the chosen use level (see equation 16). Because the optimal fertilizer use rates were calculated using the prices from profitability scenario five (i.e., acquisition price of nitrogen and maize price specific to the household), I use both those prices in this calculation as well. Table 26 shows the revenue added from fertilizer application. These values represent changes in total household income level as a result of fertilizer use at the levels observed by farmers and at calculated optimal application rates under both MVCR=1 and MVCR=2 (see Table A.8 of Appendix 8). Given this data does not provide purchase prices of maize over the whole panel, I am unable to observe longer term trends in absolute profitability using the maize prices specific to the household.

Table 26: Revenue added from the application of nitrogen (2010 prices, KSH)

District	Soil group	Actual use rates			Optimal use rates (MVCR=1)			Optimal use rates (MVCR=2)		
		2004	2007	2010	2004	2007	2010	2004	2007	2010
Machakos	3	8683	8810	17128	16944	16022	28582	16023	15042	27211
Makueni	3	10828	8904	17666	19672	16324	21688	19120	15607	21079
Meru	1	9461	5027	5332	18385	7635	7613	12542	1094	600
Mwingi	2	39764	19672	28678	56802	37536	34122	55742	36522	33556
Mwingi	3	3379	16307	30541	13429	25755	43233	12358	24702	42667
Kisii	2	13799	5530	9157	21196	8149	12978	15154	629	6524
Kisii	4	7963	4127	7641	14027	5849	9778	8679	279	2931
Siaya	3	6605	2981	7944	9284	4659	13206	6608	2239	11752
Siaya	4	11107	7755	12114	14997	9998	21738	11894	7084	20008
Bungoma	2	13398	7829	8956	22582	9186	11654	16053	1860	4669
Bungoma	3	3019	-1584	-215	4623	436	1215	45	0	0
Bungoma	4	12859	4285	4833	16972	5426	5876	9926	201	246
Kakamega	2	14272	7086	9065	16228	8151	9943	7498	1040	3241
Kakamega	3	2441	-968	1134	5129	1626	2696	249	0	0
Kakamega	4	4730	1778	2409	6534	2521	3220	763	194	0
Vihiga	3	781	-3874	-2500	1849	49	291	51	0	0
Vihiga	4	5028	-1065	2333	7540	175	2679	1671	0	0
Muranga	1	8727	4162	12394	19622	11671	20046	13301	4840	14712
Muranga	4	5800	5700	5970	22487	16402	17546	15178	7710	10957
Nyeri	1	14854	7320	7638	23211	13916	13975	16497	6544	7120
Nyeri	2	14184	9513	14109	33231	21839	27421	23073	12042	20611
Bomet	1	4180	211	907	5144	515	1024	596	0	0
Nakuru	1	-699	-2036	-1815	422	35	156	0	0	0
Nakuru	2	2442	-2021	107	2885	3	279	17	0	0
Nakuru	3	1903	-963	123	2245	5	758	162	0	0
Narok	1	-1618	-1976	-2653	0	0	0	0	0	0
Trans Nz.	4	6528	5156	7182	7760	6195	7958	1584	1290	3138
Uasin Gis.	1	2748	-1833	-731	3591	691	1102	561	0	0
Uasin Gis.	2	10610	3808	3460	13341	5062	4969	7057	450	507
Total sample		7559	2823	5596	12081	5271	8191	7374	1857	4388

Note: See equation 16 for calculations. Prices of maize and nitrogen are based on profitability scenario five (see Table 15). See Table A.8 for estimated optimal fertilizer use rates.

The negative revenue values observed in some areas and years occur when the maize yield values under fertilizer and non-fertilizer scenarios are very similar and fertilizer expense higher than the additional revenue from the small increase in maize output. As with the rest of this analysis, standard errors and deviations are very high; the sample averaged coefficient of variation on the actual use levels is around 140 percent, with tremendous variation across space and time. As such, these values should be interpreted as averages and indicators of trends, not absolute. Even so, one important finding is the huge changes in revenues between years, even when fertilizer use levels remain relatively the same. For instance, a comparison of revenues from actual fertilizer use levels in 2004 and 2007 shows that, in many places, revenues were cut in half in 2007 and sometimes negative due to high fertilizer prices. The relative measures of profitability show that 2007 was a relatively less profitable year, but these absolute profitability measures show a much more drastic picture of how those prices affected overall revenues. Notice, too, that estimated optimal levels of fertilizer use computed using relative nitrogen to maize prices can actually lead to negative revenue values where absolute nitrogen and maize values are used.

Comparing these measures to the rates of application values in Appendix 8 further illuminates the differences between relative and absolute profitability measures. In the lowlands, this table shows that there are still huge revenue gains to increasing fertilizer use to estimated optimal levels. Recall, however, that because most households in these areas are net buyers of maize, maize output is valued at the generally higher level of maize purchasing prices, which translates into relatively higher "revenue" values. In the higher potential areas, where households sometimes applied more than the estimated optimal level of fertilizer use, this table shows how revenue could improve by reducing fertilizer application rates. Furthermore, gains to changing fertilizer application rates are not nearly as large as they are in the lowlands areas, further evidence that households are applying near optimal rates already.

Chapter 7: Factors Affecting the Fertilizer Use Decision

In the last chapter, I describe fertilizer profitability and optimal use values in an unconstrained environment. However, where fertilizer is found to be profitable, I rarely observe all households fertilizing their maize fields in any amount. This suggests that farmers are operating in constrained environments. In the next two sections, I explore what the constraints to fertilizer use might be. Section 1 qualitatively explores responses from households; section 2 further investigates the decision to use commercially purchased fertilizer quantitatively using a binary response model. Given high standard deviations in estimating optimal use rates and observed use levels, this analysis will focus on the dichotomous decision to use fertilizer, not the decision to fertilize at a particular rate.

1. Qualitative Analysis of Fertilizer Use Decision

In the 2007 and 2010 surveys, households that did not use fertilizer on maize fields were asked to provide a reason for that decision. Responses to this question are found in Appendix 9. As a comparison, I include a separate table for the households found in villages not used in production function estimation because (1) the zone or soil group was deemed inhospitable to fertilizer application and/or (2) not enough households use fertilizer in these areas to predict maize response to fertilizer use. From the areas included in the production function, there are two predominant camps: (1) those that do not have cash during the necessary time frame to purchase fertilizer or deem fertilizer too expensive and (2) those that think they do not need to use fertilizer. Interestingly, all of these are demand-side reasons. In fact, only one households in 2007 (Eastern Province) gave a supply-side reason: no fertilizer was available. In the first camp, it seems that these households would use fertilizer if they had more cash or credit (an issue of latent demand); in the second camp, the households do not appear to want to use fertilizer because they think it is unnecessary for whatever objective function they seek to fulfill. These responses create an interesting picture of household perceptions of fertilizer use and profitability. While I find that fertilizer prices are generally at profitable levels, some households perceive them to be too high. Others, though, just do not have the cash available at the time necessary to purchase nitrogen, signaling a cash flow problem and the presence of credit constraints. Other less frequent responses point to other concerns about profitability (i.e., maize prices too low), the presence of information constraints (i.e., lack of advice), and the belief that fertilizer has a negative effect on the surrounding environment or soils (i.e., scorching effect).

Interestingly, the responses from this set of households are not entirely dissimilar from the responses provided by households in the villages excluded from the production function and where environmental conditions are quite different. In areas where agro-ecological conditions likely limit maize growth and the need for fertilizer (i.e., very poor soil and very low rainfall levels), a large number of households still reported not using fertilizer on maize because they did not have adequate cash or because they found fertilizer to be too expensive, implying that they would use fertilizer if it was available to them. Given similarities in responses between these very different groups of households, this calls into question how well farmers understand the conditions necessary for maize response to fertilizer.

Through this exercise, I learn that households overwhelmingly feel cash and credit constrained and infer that they would otherwise be using fertilizer if it was not for these constraints. Others feel that they "do not need to use" fertilizer, meaning they might not have the same objective function as the other group (i.e., satisfying household maize demand instead of maximizing profits).

2. Binary Response Model of Fertilizer Use Decision

In this section, I use what was learned from the qualitative analysis to further investigate characteristics of households and their operating environment which might influence their decision to use commercially purchased fertilizer on maize fields. Using these responses and a review of the literature, I populate a probit and logit model to isolate reasons for not purchasing and using fertilizer on maize fields. Unlike other studies, particularly in the technology adoption literature, I limit my sample to only those households where fertilizer use is generally profitable. I do this by taking the average AVCR value (scenario five) across the last four survey years and drop observations where the average AVCR<1. This method allows for some variability in AVCR across years and focuses on average levels of profitability. Furthermore, only the last four survey years are included (1997 excluded) due to data limitations. Refer back to Table 2 for a distribution of this sample and how it compares to the sample used in the production function.

Given the incidence of fertilizer subsidy programs in the 2010 data, I focus on the decision to use commercially purchased fertilizer, not fertilizer subsidized by the government or an NGO. This only affects observations in 2010 and very few households, at that. In this sample, 101 households claimed to receive some sort of fertilizer subsidy in 2010, but only 60 did not purchase any commercial fertilizer as a result. Furthermore, only about 23 percent of maize fields in the last four survey years went unfertilized in profitable areas with ranges from about 80 percent in the lowlands of Nyanza, 56 percent in the Eastern Lowlands, and 8 percent in the high potential Rift Valley. The variables included here seek to capture constraints related to household demographics, the size of agricultural operations, relative accessibility to markets and information, and the market environment related to the decision not to use commercially purchased fertilizer on maize fields.

Table 27: Variables used in the binary response model

Category	Variable	Measure	Level of variation
Socio-economic demographics	Age of household head (age)	Human capital constraint	Household, year
	Education of household head (educ)	Human capital constraint	Household, year
	Sex of household head (sex)	Supply constraint	Household, year
Size of farming operations and other inputs	Farm size (fsize)	Size of agricultural operations	Household, year
	Own land (tenure)	Investment in land	Household, year
	Use manure or compost (manure)	Potential substitute	Field, household, year
	Use hybrid seed (hybrid)	Potential complement	Field, household, year
Wealth and assess to credit	Asset wealth (asset)	Demand constraint	Household
	Successfully received credit (credit)	Credit constraint	Household, year
Access to fertilizer	Distance to fert. seller (dfert)	Supply constraint	Village, year
	Market price of nitrogen to maize sell price ratio (Nmaize)	Profitability expectations	District, year
Access to information	Distance to extension service (dexten)	Information constraint	Village, year
	Part of a cooperative or group (coop)	Information constraint	Household, year
	Own cell phone (cellphone)	Information constraint	Household, year
2010: Fertilizer subsidies	Received government fertilizer subsidy (subsidy)	Market shock	Household, year (2010)
2010: Post-election violence	Hh indirectly affected by PEV (PEVindirect)	Market shock	Household, year (2010)
	Hh directly affected by PEV (PEVdirect)	Market shock (household, year)	Household, year (2010)
Controls from prod. function	Soil groups (soil)	Control	Soil group
	Zone groups (zone)	Control	Zone group
	Rainfall stress (rain)	Control	Household, year
	Year (year)	Control	Year

Note: See Appendix 10 for summary statistics.

2.1. Description of variables

In this section, I discuss the variables used in the probit model. Table 27 includes a complete list of those included and what they measure. Table A.12 of Appendix 10 includes the average and standard deviation of each of these variables split by fertilizer users and non-users.

Many of these variables are included in the plethora of technology adoption studies that exist. Feder et al. (1985) provide a review of many pieces of theoretical work and empirical studies and show the significance of a large number of household-specific variables in the technology adoption process including socio-economic variables, farm size, credit constraints and human capital. Building on this review, Feder and Umali (1993) find that factors that were originally critical in the initial phases of adoption are insignificance in later stages of the diffusion cycle. I include the variables of interest in studies from across a range of countries in addition to some that appear important in the Kenyan context.

2.1.1. Socio-economic variables

There is good reason to believe that the socio-economic status of the household influences its decision to use fertilizer and, in particular, to make decisions about fertilizer use congruent with profitability. Here, I include the age and education of the household head as a proxy for human capital, experience, and the likelihood of making profitable input decisions. A large number of studies have empirically verified the link between education and the early adoption of new technologies (see Feder et al. 1985 for a review). While fertilizer is by no means a new technology in Kenya, making the decision to use fertilizer when and where it is profitable is likely to be correlated with education in the same way. For example, Huffman (1974) find that corn farmers in the Midwestern United States with higher education levels make more “allocatively efficient” fertilizer use decisions in dynamic profitability environments, much like they appear to be in Kenya. The number of years of formal education of the household head is included, as well as the age of the household head, a proxy for experience level and a potential substitute for formal education.

A lot of literature investigates the differences between female and male-headed households and how those differences contribute to on-farm decisions and technical efficiency. In a study by Doss and Morris (2001) about the adoption of a range of inputs in Ghana, they found that women are less likely to have access to complementary inputs of a technology (i.e., land, labor, extension services), resulting in lesser use of the technology in question. Similarly, Doss (2001) describes the complexity and heterogeneity of sex and gender dynamics in African households, making the point that generalizations are quite difficult. Doss and Morris (2001) also distinguish between the sex of the household head and the sex of the farmer, noting that this specification leads to different results. However, about 85 percent of households in this sample claim that the head of the household makes decisions with respect to the farm. As such, the sex of the household head, described with a dummy variable, is used here.

2.1.2. Size of farming operations and other inputs

While all households included in this analysis have maize fields, the size and intensity of operations may differ substantially. In their review, Feder et al. (1985) show that farm size is generally a significant determinant of adoption of lumpy technologies (e.g., irrigation equipment or tractor) but not necessarily divisible inputs like fertilizer (i.e., the farmer can decide to use 1 kg or 100 kgs of fertilizer). In fact, some studies show that larger farms are more likely to use fertilizer while others show that smaller farms have the advantage. Here, I attempt to measure the effect of size and scale of farming operations by including farm size (in hectares).

Not only does size of agricultural operations matter, but also ownership over them. Referring to lumpy and indivisible technology investments, Gebremedhin and Swinton (2003) find that having secure rights to land in Ethiopia created incentives for farmers to invest in longer-term soil conservation techniques. While fertilizer is a divisible input, farmers may associate using it with longer-term plans for maintaining soil nutrients and land productivity. Li et al. (1998), for example, find that farmers in China using private land were more likely to use higher levels of fertilizer than those farming on collective land. Given these observations, I include a dummy variable for households that own their land with a title deed in the model.

Another important aspect of the fertilizer use decision is the other inputs used alongside or in place of it. Manure, for example, may be perceived by farmers as a substitute for fertilizer use, particularly when large quantities are available. Research by Abdoulaye and Sanders (2005) in Niger shows that farmers may also consider manure a complementary input, used alongside inorganic fertilizer in order to help hold water, especially in sandy soils. Waithaka et al. (2007) investigate manure and fertilizer use in the Vihiga district of Kenya and find them positively correlated suggesting, again, that farmers regard the two as complementary. In order to test whether farmers perceive manure and fertilizer as complements or substitutes and how these perceptions frame the fertilizer use decision, I include a dummy variable in the model where the household applied manure or compost to their maize fields.

Hybrid seeds, too, may be considered a complementary input. In Swaziland, Rauniyar and Goode (1992) show that high-yielding seed varieties are most often adopted in a "package" with inorganic fertilizer. Given the very high correlation between hybrid seed and fertilizer use, I also include a variable to denote which fields have new hybrid maize seeds. This could be one possible explanation as to why so many households in rain stressed areas (the lowlands) have started using fertilizer later; hybrid seed use, which improves the response of fertilizer, is not as profitable there. I test this claim by adding a dummy variable to the model where the household used new hybrid maize seed.

2.1.3. Access: cash, credit, fertilizer markets, information

There are several constraints to access that might limit farmers' ability to procure or use fertilizer including access to cash, credit, fertilizer markets and information. Because available income and, in particular, the flow of available income over the year, are difficult to accurately specify for households, I use household asset wealth (averaged over time) as an indicator of financial liquidity and purchasing power. Where income and assets are limited, households are likely to need credit in order to purchase inputs. A large number of empirical studies show the importance of credit constraints in limiting fertilizer use (for example Coady 1995 on Pakistan; Croppenstedt et al. 2003 on Ethiopia; Odhiambo and Magandini 2008 on South Africa). While I do not necessarily know which households are constrained by credit, I do know which households sought credit and were successful, the group opposite that of interest. I denote these households using a dummy variable.

Physical access to the fertilizer market or dealer is also of interest. While practically no households report distance to markets (or supply side constraints, in general) as the most important factor limiting their ability or incentive to procure fertilizer (see Appendix 9), I include the village-averaged kilometers from the household to the nearest fertilizer seller in order to measure its effect. While households do not cite accessibility as a constraint to use, the transportation cost is a component of the AVCR metric used in the profitability scenario used to create this sample. I also include the relative price of nitrogen (market price) to the selling price of maize given the large number of households claiming that the price of fertilizer was prohibitively high.

Lastly, access to information on proper application of fertilizer use is a cited deterrent by households. I measure access to information using three possible forms of information transmission: (1) extension services, (2) cooperatives or other formal groups, and (3) mobile phones or land lines. Previous studies have found mixed results on the utility of government extension service in the use of fertilizer (e.g., Freeman and Omiti 2003 in Kenya; Kaliba et al. 2000 in Tanzania) and that extension agents are likely to suggest blanket fertilizer use recommendations irrespective of farmer or geography specific conditions (e.g., Snapp et al. 2003). I test for the significance of access to extension service in making fertilizer use decisions in line with profitability measures by including the distance to the nearest extension agent in the model. Others have noted the importance of social capital in making good fertilizer use decisions (e.g., Isham 2002 in Tanzania). I include a dummy variable for households that are members of a cooperative or other formal group as a proxy for social capital and the associated information flows. Finally, with the proliferation of mobile phones across Kenya (see *The Economist* September 26th 2009 issue), I test whether or not owning a phone (either mobile or land line), and therefore a means of accessing remote information, encourages farmers to make better decisions about fertilizer use by including a dummy variable for households that own either type of phone.

2.1.4. Targeting of government fertilizer subsidies (2010 only)

Following the successes of Malawi and with a pledge at the African Fertilizer Summit in Nigeria, a proposal was developed in 2006 by the Ministry of Agriculture in Kenya for a multi-million dollar improved seed and fertilizer subsidy program, the National Accelerated Agricultural Inputs Access Program (NAAIAP), aimed at reaching 2.5 million farmers. The main features of the program were to provide farmers with less than 2.5 acres of land basic inputs to cover at least one acre of land through a voucher redeemable at a local retailer. These characteristics are similar to other "smart" subsidy programs rolling out across Africa, aimed at building on already established private sector networks and targeting those households that would otherwise be unable to purchase the inputs (see Dorward 2009; Minot and Benson 2009; Banful 2011 for more on these subsidy programs). Donors, however, were tepid on supporting the efforts given perceptions that the program was too large, too expensive, and scaled up too quickly without the existing capacity necessary to do so. In the absence of donor support, the Government of Kenya was able to pay for only a portion of the first year of the program, meaning the originally intended project required substantial downward revision.

Table 28: Frequency of government fertilizer subsidy recipients by district

District	NAAIAP	Other gov't subsidy
Machakos	0	1
Makueni	34	1
Meru	2	3
Kisii	3	6
Kisumu	0	2
Siaya	0	12
Bungoma	6	4
Kakamega	3	1
Muranga	8	1
Nyeri	11	4
Bomet	0	2
Nakuru	4	9
Narok	1	0
Trans Nzoia	5	14
Uasin Gishu	11	14
Total	88	69

Note: This table includes all households in the 2010 survey, not just those used in the production function and profitability analysis.

The final wave of this dataset (2010 survey) shows which households received the NAAIAP subsidy in any year between 2007 and 2010. Representative of the significantly smaller program, only 85 of the 1243 households in the full panel received assistance under NAAIAP. The government also had other fertilizer subsidy programs occurring simultaneously, but more focused on larger farmers and surplus areas. Table 28 shows the distribution of households by subsidy type across the entire sample (not just those in the profitability analysis). I test the claim that receiving a subsidy was a significant determinant of (or deterrent to) using commercially purchased fertilizer in 2010 by including a dummy variable to denote which households received any government fertilizer subsidy.

2.1.5. Post-election violence of 2007-2008 (2010 only)

The disputed presidential election of December 2007 produced widespread violence and upheaval throughout the country. Over the month of January 2008, official figures state that over 1,200 people were killed, many more injured, 300,000 displaced from their communities, and that property destruction was widespread, including the burning of about 50,000 houses (UNHCR 2008). The food and agricultural system effects of the violence were said to be vast. Given that much of the violence took place in surplus maize production areas (the Rift Valley, Western and Nyanza provinces), concerns over reduced yields, disrupted input and output markets and heightened food insecurity in the coming agricultural seasons were pervasive. The violence ended in late February 2008, but tensions remained high, agricultural marketing channels disrupted, and many households still unable to return to their homes and farms.

Ideally, I would include a subjective measure of the intensity of the violence in a particular location to test if and how the post-election violence affected fertilizer use in the 2010 season. However, most of these statistics are reported at the district level, making collinearity a problem. Furthermore, most of the violence was concentrated in the net surplus areas where fertilizer use is widespread. Because violence was not randomly allocated throughout the country, it will be difficult to retrieve a clean estimate of how the violence affected fertilizer use. For reference, though, see Table 29 for the number of deaths attributed to the post-election violence by district and Figure A.6 of Appendix 12 for a map of internally displaced persons.

Instead, then, I include two variables in the model. In the 2010 survey, households were asked whether or not they were affected by the violence and, if so, was it directly[15] or indirectly[16]. I include each of these self-reported claims as a dummy variable in the model.

Table 29: Number of deaths attributed to the 2007-2008 post-election violence

District	Deaths
Bomet	4
Bungoma	7
Kakamega	31
Kisii	9
Kisumu	81
Nakuru	213
Narok	19
Siaya	10
Trans Nzoia	104
Uasin Gishu	230
Vihiga	18

Source: WAKI report (2008).

2.2. Model specification

In this section I describe the binary response model used in regression, estimated using both a probit and logit model and, for comparison, a linear probability model (LPM). I assume that the probability Y of using commercially purchased fertilizer on a given field *i* at household *j* during year *t* where it is profitable takes the following form (see Table 27 for variable abbreviations):

$$Y_{ijt} = \alpha_1 + \beta_1 age_{jt} + \beta_2 educ_{jt} + \beta_3 sex_{jt} + \beta_4 fsize_{jt} + \beta_5 tenure_{ijt} + \beta_6 manure_{ijt} + \beta_7 hybrid_{ijt} + \beta_8 asset_{jt} + \beta_9 credit_{jt} + \beta_{10} Nmaize_{jt} + \beta_{11} dfert_{jt} + \beta_{12} dexten_{jt} + \beta_{13} coop_{jt} + \beta_{14} cellphone_{jt} + \beta_{15} subsidy_{jt} + \beta_{16} PEVdirect_{jt} + \beta_{17} PEVindirect_{jt} + \beta_{18} soil + \beta_{19} zone + \beta_{20} rain + \beta_{21} year + \mu \quad (22)$$

[15] Direct effects include: household members displaced, lost family member, injury of household member, property destroyed or lost, crops destroyed, and lost livestock.

[16] Indirect effects include: hosted or supported internally displaced persons (IDPs), disruption of produce markets, high commodity prices, disruption of schooling, general insecurity, disruption of transport, heightened land insecurity, farming delayed, and relative's property destroyed.

As described in Table 27, the variables included in this model are at a number of different levels of aggregation. For example, the manure and hybrid variables describe specific field-level features while age and cell phone are household specific. Furthermore, some variables are observed (or averaged) at the village level (i.e., distance to the nearest fertilizer dealer) or district level (i.e., relative price of nitrogen to maize).

In order to control for whatever variation in fertilizer choice remains beyond those variables that I can accurately specify, I include several variables from the production function to absorb the variation. I add the three variables used as interactions with the nitrogen variable in the production function (i.e., zone groups, soil groups, rain stress). Then, as a final control, I absorb variation specific to time using a year dummy variable. If it is the case that there are certain constraints or restrictions on fertilizer use specific to a given area or year beyond what I am able to characterize in the model, these variables should pick up on those remaining average characteristics. This method, though, does introduce a lot of collinearity into the model (condition score of about 40). Even with a lot of collinearity, the variables of interest should have enough variation to produce good estimates (see Appendix 10).

2.3. Results and discussion

The results of the probit and logit regressions are found in Table A.13 of Appendix 11. The interpretation of coefficients from these non-linear models is slightly more conceptually difficult given the functional form. For this reason, I include the partial effects in a separate table (Table A.14) where I add the coefficient estimates from the LPM. The partial effects estimates are used in the discussion that follows. Appendix 11 shows a great deal of similarity in estimates among the three model types, as expected. I focus on the probit model estimates here.

In general, this model has great predictive power (85 percent of cases are properly predicted); however, disaggregating by zone shows some variation in goodness of fit across space. The model accurately predicts only about half of the cases where fertilizer was used in Siaya where only about 30 percent of fields were fertilized while closer to 95 percent in the high potential and highland areas where over 90 percent of fields were fertilized. Of non-users, the model predicts about 70 percent in the Eastern Lowlands and 97 percent in the Western Lowlands. Non-users in the high potential and highlands areas are rarely accurately predicted. For example, only 16 of the 109 non-fertilized fields in the High Potential Maize zone are correctly classified; however, non-fertilized fields make up only about 8 percent of maize fields in this zone across the last four survey years. This means the model does better job of predicting fertilizer non-use in the areas that use it less, not the areas with already very high percentages of fertilized fields.

First, I look at the variables attempting to measure those reasons provided by households in the qualitative section to test for the overall quantitative significance of their claims. For those many households saying they did not have enough cash or cash at the necessary time, I measure this using an average of household asset wealth which, in this model, produces a very small effect and is not significant. This could mean that assets are not directly correlated with cash on hand or that asset wealth is not a good indicator of the liquidity of a household. Perhaps assets serve as long term savings for the household, whereas cash, the resource they claim to lack, is still scare in households with a large amount of accumulated assets. I gain further insight in to the question of cash constraints by looking at the dummy variable for households that received credit (of any type) in a given agricultural season. Again, this variable is insignificant, showing that households that received credit are not significantly more likely to use fertilizer on maize fields than those that do not gain credit. While not a perfect measure of credit constraint, it does show that the two are not well correlated. Another claim was that fertilizer was too expensive. For this, I look at the relative price of nitrogen (market price) to maize (selling price), which is significant at the 99th percent confidence level. All else equal, a 1 KSH increase in the price of nitrogen relative to maize makes a household 1 percent less likely to use commercial fertilizer. This shows that the gap between nitrogen prices and maize prices is considerably different across users and non-users of fertilizer, as would be expected given the responses from households.

Almost all of the variables representing fertilizer market access (i.e., distance to fertilizer seller) and information access (i.e., part of a cooperative or group, own a cell phone) are significant and exhibit the expected signs. For example, for every 1 kilometer a household is away from a fertilizer retailer makes them about 1 percent less likely to use fertilizer. On the positive side, households that are members of a cooperative or group are 6 percent more likely to use fertilizer while those owning a cell phone are 4 percent more likely. The significance of most of these numbers suggests that being in a more remote location without access to markets or advice does have a negative effect on fertilizer use, even where it is profitable to use it. Moreover, having access to information positively contributes to fertilizer use.

There are a number of other household socio-economic variables and characteristics of the farm that are useful to explore. In terms of socio-economics, only the age of the household head is not a predictor of commercial fertilizer use. An additional year of formal education makes the household 0.7 percent more likely to use fertilizer. Female headed households are 3 percent less likely to use commercial fertilizer. Furthermore, of the variables characterizing the size of farming operations, the size of the farm is a significant determinant of fertilizer use but land ownership is not. A 1 hectare increase in farm size makes a household 0.8 percent more likely to use commercial fertilizer.

The other inputs used on maize fields are quite telling. Those using manure on their maize fields in profitable areas are 13 percent less likely to use fertilizer while those using new hybrid seed are about 23 percent more likely to use fertilizer. These findings provide insight into how farmers view the relationship between inputs; manure may function as a substitute while hybrid seeds a complement. It is unclear to what extent farmers decide to use manure and compost because they are unable to afford inorganic fertilizer or if they make the decision to use manure irrespective of price. Only a handful of households mentioned the desire to farm organically as a reason for not using inorganic fertilizer; however, access to and use of manure could also have been part of the common "no need to use" response. The effect of hybrid seed use is not surprising given the two are recommended as an input package.

Receiving the NAAIAP or other government subsidy somewhere between 2007 and 2010 did make households about 35 percent less likely to use commercial fertilizer in profitable areas in 2010. This value provides insight into the potential displacement of commercial fertilizer from rural areas when fertilizer subsidies are introduced. Further work on fertilizer demand would help to illuminate the "crowding out" effect of the various government fertilizer subsidy programs happening concurrently in Kenya (see Ricker-Gilbert et al. 2011 on Malawi; Xu et al. 2009 on Zambia). With respect to post-election violence, neither of the household specific claims are significant in determining use. This could be for a number of different reasons: (1) the violence happened to be concentrated in more agricultural productive areas, (2) there is measurement error in the self-reported claims, or (3) that the effects of the violence had dissipated by the 2010 main season. Note, also, that the 2010 dummy variable is the only one that is not positive and statistically significant, meaning households in 2010 were not as likely as those in 2004 and 2010 to use commercial fertilizer, representing a set back in profitable fertilizer use expansion in this survey year.

It should be noted, however, that my model essentially has a double error term, one carried over from the production function (i.e., households could be placed in the wrong category given the error term in the production function) and one from the probit model itself. As such, I cannot make overwhelming conclusions about constraints to commercial fertilizer use. Instead, I can note that over a continuum of space where fertilizer use is estimated to be profitable, there are several variables that appear significant in the fertilizer use decision, namely the use of other inputs, the education and sex of the household head, the distance to fertilizer sellers, and a range of information accessibility variables. Moreover, the non-use of fertilizer is concentrated in a few areas of the country, meaning these results are more likely to mimic constraints in those areas than all of Kenya.

Chapter 8: Summary and Conclusions

1. Summary

Based on experimental trials, it is widely perceived that Kenyan farmers are under-utilizing chemical fertilizer and that tremendous gains in maize output could be realized through the continued promotion of fertilizer use. However, very little evidence from farmers' fields, based on the constraints they face, exists to back this claim. For years, researchers have noted increases in national level fertilizer consumption levels and a gradual reduction in fertilizer prices in agriculturally productive areas since market liberalization in the mid 1990s. This thesis set out to provide a more in-depth and researched picture of fertilizer profitability and use patterns on maize over time and across Kenya using data from five rounds of a nationally representative panel household survey. In general, I find that the gap between where it is profitable to use fertilizer and where households use it is not nearly as large as one might expect. While there is likely room for expansion in some of the lowlands areas, households in the most agriculturally productive areas are using fertilizer at or beyond the most profitable levels.

I estimate maize yield response using a modified quadratic production function, controlling for unobserved household heterogeneity through correlated random effects (CRE) and with careful consideration of the diverse biological and ecological environments available to Kenyan farmers. With these regression estimates, I calculate the marginal physical product (MPP) and average physical product (APP) of nitrogen by district, soil group, and year, finding considerable differences across the country, particularly between areas of low and high agricultural potential. A range of other inputs, including manure and hybrid seeds, also contribute positively to maize output, while others, like rain stress, lead to reduced yields.

Using marginal and average products of nitrogen and five relative price scenarios, I estimate a range of fertilizer profitability measures using marginal value cost ratios (MVCRs) and average value cost ratios (AVCRs). I use the standard market price of nitrogen then add a transport cost using observed distance between households and fertilizer retailers to create an "acquisition cost" of fertilizer. On the output side, I use estimates of both expected selling and buying prices of maize given a large number of net maize buying households across Kenya. I find that MVCR or AVCR values vary considerably with the relative price chosen; however, never does a change in relative price scenario push the overall profitability level from profitable to unprofitable in a given district and year. When looking at fertilizer use patterns alongside calculated profitability levels, I find a large number of households using fertilizer where MVCR values are below two, suggesting that the "MVCR=2 rule" used throughout the literature is not necessarily appropriate for more mature input markets and where learning is taking place.

When assessing fertilizer profitability measures alongside actual fertilizer application rates over time, I find a closing gap between where it is profitable and where households are using, however with considerable variation across space. In the Eastern and Western Lowlands, households have significantly increased fertilizer use over time, both the percentage using and the amount they apply, but could increase income further by using slightly more. In the Western Province and Rift Valley, however, households are using either at optimal levels or slightly beyond. In this case, households either need to cut back on levels or consider applying nitrogen fertilizer in conjunction with lime in order to reduce soil acidity and ensure fertilizer is a long term profitable strategy. In the highlands areas of western and central Kenya, a lack of concavity in the production function creates high profitability measures, which should mean that expansion of fertilizer use would be profitable. Given a high incidence of fertilizer use and the high volumes already used, these findings should be interpreted cautiously.

I estimate optimal nitrogen application levels by district and soil group using two risk scenarios: where the MVCR=1 and MVCR=2. In general, these findings further corroborate earlier results. By 2010, households in the lowlands areas had approached profitable levels where MVCR=2 but can expand current use levels by 5-10 kg/ha in order to reach the most profitable levels (where MVCR=1). Households in the high potential and transitional areas are using at or beyond estimated optimal levels. Even here, there has been a slight increase in use rates since 1997, not all of which appears profitable. In the highlands, the model produces unreasonable estimates to optimal use rates, again, due to lack of concavity in the production function.

In addition to MVCRs and AVCRs, both measures of relative profitability, I also compute two measures of absolute profitability, the net gain from the last unit of fertilizer and total revenue added from fertilizer application. While relative nitrogen to maize prices do not vary considerably over time, changes in absolute prices means that total revenue from fertilizer use varied much more substantially between years. For example, higher fertilizer prices in 2007 meant farmers' revenues, even where application rates remained unchanged, were cut in half compared to 2004 levels. Furthermore, while optimal nitrogen application rates are based on relative profitability measures, higher revenues are realized when considering absolute prices.

When asked why they do not use fertilizer on maize, households overwhelming say they are either cash constrained or do not need to use fertilizer. These, in addition to other household responses and a review of the literature, inform the creation of a binary response model for commercial fertilizer use. By confining my sample to only those areas where fertilizer use is profitable and to the final four survey years, I attempt to isolate the constraints on households limiting an otherwise profitable fertilizer use decision. While only about 23 percent of the fields in this sample are not fertilized using commercial fertilizer, I learn that distance to the nearest fertilizer seller (despite its drop over time), the ratio of nitrogen to maize price, a range of information accessibility variables (i.e., own a cell phone, member of a cooperative or grower group), the choice of other inputs (i.e., manure and hybrid seeds), and education and sex of the household head are significant determinants of the fertilizer use decision where profitable. Furthermore, receiving a government fertilizer subsidy somewhere between 2007 and 2010 made households 35 percent less likely to use commercial fertilizer on maize fields. Further research on commercial fertilizer displacement is needed to better understand the market effects of recent fertilizer subsidy efforts.

2. Limitations

The incidence of high standard deviations in input and output levels coupled with high standard errors in some model coefficients highlights the importance of considering local level variation and conditions when analyzing input response. I attempt to control for this important variation where the data is available but believe this analysis would be improved with even more local level information (e.g., more specific soil data). I capture the over-arching differences between locations, but further disaggregation might lead to better optimal fertilizer use recommendations. There are a number of key variables that would help to untangle the differences in fertilizer response over a limited geographic area that I am unable to control for (e.g., timing of planting, timing of fertilizer treatments, more specific seed type, slope of field, average main season temperature, pests). A large sample size helps to provide better estimates, however these omitted variables would help to capture some of the variation.

Furthermore, in order to reduce measurement error in prices, I average to the district level, which reduces some of the important variability in prices that might contribute to household level fertilizer use profitability. Some households may receive better prices than others given social connections or market knowledge. Not only that, but I also use one price per season whereas it is well-known that both input and output prices can fluctuate substantially over a year, again causing some lack of precision in the profitability metrics. What I create here is a well-approximated picture of average fertilizer response using a large sample size and extended time period. While this analysis goes far beyond that has been attempted by other researchers, there is still considerable room for improvement in order to better understand the complexities of fertilizer use profitability and farmers' decisions to use the input.

3. Conclusions

This study makes a number of contributions, both in its approach and policy-relevant findings. First of all, I find that maize response to fertilizer can vary considerably across space and time within the same country, meaning recommended fertilizer use levels should vary accordingly. Furthermore, this analysis is the first using household data to look at maize response to fertilizer application in the lowlands and eastern areas of the county where fertilizer use has been increasing with time. Secondly, I construct a number of relative price scenarios where maize market standing of the household and distance to the nearest fertilizer retailer are considered. Perhaps the first example of a carefully delineated comparison between actual relative prices, I find that the chosen price of maize (e.g., buying or selling prices) and including the transport cost of fertilizer does change the level of profitability but never substantially, particularly in recent years. Third, I find many households using fertilizer where the MVCR is much less then 2. This calls into question the necessity of the "MVCR=2 rule" used throughout the literature, particularly where fertilizer has been available for many years and where prices are well specified. Fourth, I show how absolute and relative profitability measures can produce different results. Given falling prices of nitrogen and maize over time, understanding both the relative profitability, which describes incentives to use nitrogen, and absolute profitability, which describes changes in actual household revenue, is important for input decision making.

While cognizant of the fact that households and field level specifics should be considered when making decisions on fertilizer use, I find that, in general, tremendous expansion of fertilizer use in Kenya is not necessarily a profitable strategy. These findings provide counter-evidence to claims that increased fertilizer use is critically important to expanding maize output, improving food security, and helping to lift farmers out of poverty. In fact, in some high productive areas, households may actually be over-using fertilizer, a claim which requires further corroboration. In other areas of the country, however, there are still some constraints to fertilizer use where it is profitable to do so, particularly in the lowlands areas.

Evidence of a closing gap, only about 20 percent of maize fields in areas where fertilizer use is profitable have gone unfertilized since 2000. Helping farmers to access cash or credit or helping them reach an appreciable maize surplus to break out of the cycle of net buyer status might enable them to see the household income gains associated with using inorganic fertilizer at profitable levels. What would make fertilizer use more economically profitable for all households would be a continued reduction in fertilizer prices and transportation costs. Similarly, stable maize selling and buying prices would help farmers make input decisions in line with more accurate expected profitability calculations. The risks associated with not producing enough maize to feed a household, apparent from the large number of households shifting between net buyer and net seller status, makes planning for the agricultural season tremendously difficult and risky for small scale farmers. Providing an institutional environment which promotes ease of access to inputs, markets, and reliable prices will most likely translate into increased incomes for maize producing households and a more efficient and productive agricultural sector in Kenya.

APPENDICES

Appendix 1: Example computations of Liu and Myers yield index by field composition

In this Appendix, I provide more detail on how the Liu and Myers yield index works under different field compositions and relative price scenarios that frequently occur in the data set used in this analysis. Because I do not observe how spatially the field was planted or what portion of it was devoted to which crop, I use observed output values and transform them into maize equivalents using relative prices. Below, I use six examples to show how the Liu and Myers yield index produces output as "maize equivalents" and which field types are kept for use in analysis. Note how the output value is some combination of "kilograms of yield" and "revenue," depending on the nature of the field. Refer to equation 18 on page 39 and related discussion for more details.

Example 1: All maize field (monocropped)

Consider a monocropped field that yielded 5000 kgs of maize grain per hectare. The maize grain selling price in the district was 25 KSH/kg. Using the Liu and Myers index, the output on this field is 5000. For all monocropped fields where only maize grain is harvested, yield as computed using the Liu and Myers index, is equivalent to total kilograms of maize harvested per field.

$$Y_{ijt} = \frac{\sum_n Y_{is} P_s}{P_m} = \frac{5000 \text{ kg} * 25 \text{ KSH/kg}}{25 \text{ KSH/kg}} = 5000$$

Furthermore, because maize constitutes 100 percent of the potential revenue from this field (like all monocropped fields), this type of field is always kept for use in analysis.

Example 2: Maize harvested as grain and green maize (monocropped)

Consider a monocropped maize field where the household harvests some of the maize green and some of it as grain. Because green and grain maize go for different prices on the market, the output index does not work the same as it did in Example 1. Instead, I consider the two crops separately. Suppose the household harvested 6000 kilograms of maize grain and 150 kilograms of green maize. The maize grain selling price in the district was 20 KSH/kg while the green maize selling price was 10 KSH/kg. Using the Liu and Myers index, the output on this field is 6075, meaning green maize is valued at a lower weight due to its lower output price.

$$Y_{ijt} = \frac{\sum_n Y_{is} P_s}{P_m} = \frac{(6000 \text{ kg} * 20 \text{ KSH/kg}) + (150 \text{ kg} * 10 \text{ KSH/kg})}{20 \text{ KSH/kg}} = 6075$$

Despite the fact that grain and green maize are considered separately in the yield index computation, this field is still considered monocropped and, therefore, maize still constitutes 100 percent of the potential revenue from this field. As such, this type of field is always kept for use in analysis.

Example 3: Maize and beans in alternating rows (intercropped)

Consider a maize field where beans are found in rows between maize (i.e., intercropped). This is a common field type in Kenya because households often consume maize and beans together. On this field, 2250 kgs of maize grain and 135 kgs of beans were harvested with selling prices of 30 KSH/kg and 45 KSH/kg respectively. This produces an output value of 2452.5. In this case, beans are weighted more (per unit weight) than maize grain because the beans had a higher market value than maize grain; this is opposite of Example 2 where green maize was weighted less than maize grain.

$$Y_{ijt} = \frac{\sum_n Y_{is}P_s}{P_m} = \frac{(2250 \text{ kg} * 30 \text{ KSH/kg}) + (135 \text{ kg} * 45 \text{ KSH/kg})}{30 \text{ KSH/kg}} = 2452.5$$

Before deciding to use this field, I must ensure that at least 25 percent of the potential revenue from this field came from maize. Per the below calculation, about 92 percent of the revenue from this field would have come from maize. This field is kept for use in production function estimation.

$$\text{revenue from maize} = \frac{Y_m P_m}{\sum_n Y_{is}P_s} = \frac{2250 \text{ kg} * 30 \text{ KSH/kg}}{(2250 \text{ kg} * 30 \text{ KSH/kg}) + (135 \text{ kg} * 45 \text{ KSH/kg})} = 0.92$$

Example 4: Maize and cowpea rows with a guava and orange tree (intercropped)

Consider a maize field where cowpeas are found in rows between maize (i.e., intercropped) with both a guava and orange tree on one side. On this field, the household harvested 5075 kgs of maize grain, 450 kgs of cowpeas, 45 kgs of guava and 15 kgs of oranges which were valued at 40 KSH/kg, 50 KSH/kg, 12 KSH/kg and 17 KSH/kg respectively. This produces an output index of 5657. The 510 kgs of non-maize output was valued at 582 kgs given the abundance of relatively higher valued cowpeas.

$$Y_{ijt} = \frac{\sum_n Y_{is}P_s}{P_m} = \frac{(5075 \text{ kg}*40 \text{ KSH})+(450 \text{ kg}*50 \text{ KSH})+(45 \text{ kg}*12 \text{ KSH})+(15 \text{ kg}*17 \text{ KSH})}{40 \text{ KSH/kg}} = 5657$$

Before deciding to use this field, I must ensure that at least 25 percent of the potential revenue from this field came from maize. Per the below calculation, about 90 percent of the revenue from this field would have come from maize. So, while four additional crops on a maize field might seem like a lot, this field is clearly dominated by maize. This field is kept for use in production function estimation.

$$\text{revenue from maize} = \frac{Y_m P_m}{\sum_n Y_{is} P_s} = \frac{5075\text{ kg} * 40\text{ KSH/kg}}{(5075\text{kg}*40\text{KSH})+(450\text{kg}*50\text{KSH})+(45\text{kg}*12\text{KSH})+(15\text{kg}*17\text{KSH})} = 0.90$$

Most fields in this data set resemble this one (i.e., maize intercropped with beans or another legume, potentially a green vegetable or squash, and with a fruit tree or two). So long as more than 25 percent of potential field revenue is derived from maize, then the field is kept for use in analysis.

Example 5: Maize and coffee found on the same field (intercropped)

Consider a field with both maize and coffee. On this field, the household harvested 1000 kgs of maize and 1750 kgs of coffee which were valued at 25 KSH/kg and 45 KSH/kg respectively. This produces an output index of 4150. Because coffee had a higher market value, it was weighted much more heavily than the maize grain, converting a combined 2750 kilograms of output into the much higher 4150 (i.e., a difference of 1400).

$$Y_{ijt} = \frac{\sum_n Y_{is} P_s}{P_m} = \frac{(1000\text{ kg} * 25\text{ KSH/kg}) + (1750\text{ kg} * 45\text{ KSH/kg})}{25\text{ KSH/kg}} = 4150$$

Before deciding to use this field, I must ensure that at least 25 percent of the potential revenue from this field came from maize. Per the below calculation, only about 24 percent of the revenue from this field would have come from maize. So, while this field only contains maize and one other crop, maize does not constitute the dominant output. This field is not used in analysis.

$$\text{revenue from maize} = \frac{Y_m P_m}{\sum_n Y_{is} P_s} = \frac{1000\text{ kg} * 25\text{ KSH/kg}}{(1000\text{ kg} * 25\text{ KSH/kg}) + (1750\text{ kg} * 45\text{ KSH/kg})} = 0.24$$

Fields with major cash crops (i.e., tea, rice, sisal, pyrenthrum) are always dropped from analysis. Because coffee, one of the more traditional cash crops, can be a minor crop on a mostly maize field, fields with coffee are considered. In this case, the field was some mixture between a maize field and a coffee field. With even 100 kgs more of maize output, this field would have met the criterion of potential field revenue exceeding 25 percent.

Example 6: Low maize harvest relative to beans, sukuma wiki and groundnuts

Consider a field where maize and beans are found in alternating rows alongside a small patch of sukuma wiki (kale) and groundnuts. On this field, the household harvested 75 kgs of maize, 100 kgs of beans, 75 kgs of sukuma wiki and 20 kgs of groundnuts which were valued at 35 KSH/kg, 50 KSH/kg, 15 KSH/kg and 120 KSH/kg respectively. This produces an output index of 319. The beans and groundnuts had a relatively higher market value than maize while sukuma wiki had a lower value.

$$Y_{ijt} = \frac{\sum_n Y_{is} P_s}{P_m} = \frac{(75\text{ kg}*35\text{ KSH})+(100\text{ kg}*50\text{ KSH})+(75\text{ kg}*15\text{ KSH})+(30\text{ kg}*120\text{ KSH})}{35\text{ KSH/kg}} = 319$$

In this instance, it is pretty obvious that the maize crop either failed or that maize constituted only a small portion of this field. To be sure, however, I calculate the potential revenue from maize. Per the below calculation, about 24 percent of the revenue from this field would have come from maize, which almost seems high given the very low value of maize output (compared to previous examples).

$$\text{revenue from maize} = \frac{Y_m P_m}{\sum_n Y_{is} P_s} = \frac{75\text{ kg} * 35\text{ KSH/kg}}{(75\text{kg}*35\text{KSH})+(100\text{kg}*50\text{KSH})+(75\text{kg}*15\text{KSH})+(20\text{kg}*120\text{KSH})} = 0.24$$

If even an additional 10 kgs of maize had been harvested, this field would have met the criterion of at least 25 percent of revenue coming from maize. Still, though, the portion of maize on the field would have been low. I control for these situations by including the portion of revenue from maize as an explanatory variable in the production function.

Appendix 2: Percent of major nutrients in each kilogram of fertilizer type

This table shows how individual fertilizer types (i.e., what is observed in the data) are broken down into their constituent nutrient parts. These values are calculated using the ratio in the fertilizer type column where the ratio stands for the N, P_2O_5, and K_2O respectively. P constitutes 43.6 percent of P_2O_5 while K constitutes 83 percent of K_2O. For example, DAP (18:46:0) contains 18 percent nitrogen, about 20 percent (43.6*46) phosphorous and 0 percent potassium.

Table A.1 Percent of major nutrients in each fertilizer type

Fertilizer type	N	P	K
DAP (18:46:0)	18	20.06	0
MAP (11:52:0)	11	22.67	0
TSP (0:46:0)	0	20.06	0
SSP (0:22:0)	0	9.59	0
NPK (20:20:0)	20	8.72	0
NPK (17:17:0)	17	7.41	0
NPK (25:5:+5s)	25	2.18	0
CAN (26:0:0)	26	0	0
ASN (26:0:0)	26	0	0
UREA (46:0:0)	46	0	0
DSP (0:19.43:0)	0	8.47	0
SA (21:0:0)	21	0	0
NPK (23:23:23)	23	10.03	19.09
NPK (20:10:10)	20	4.36	8.30
NPK (23:23:0)	23	10.03	0
NPK (17:17:17)	17	7.41	14.11
NPK (18:14:12)	18	6.10	9.96
NPK (15:15:15)	15	6.54	12.45
NPK (14:14:20)	14	6.10	16.6
NPK (26:5:5)	26	2.18	4.15
NPK (22:6:2) + TE	22	2.62	9.96
NPK (22:11:11)	22	4.80	9.13
Foliar feeds (12:10:7)	12	4.36	5.81
Mavuno basal (10:26:10)	10	11.34	8.30
Kero green (10:46:10)	10	20.06	8.30
Mavuno top dress (30:8:6)	30	3.49	4.98

Note: For TSP, the P_2O_5 component can range from 40 to 54 percent. 46 percent is used here. For SSP, P_2O_5 component can range from 18 to 22 percent. 22 percent is used here. For rock phosphate, the P_2O_5 component can range from 2 to 4 percent. 3 percent is used here.

Appendix 3: Dealing with collinearity between phosphorous and nitrogen

In a complete quadratic production function, both the nitrogen and phosphorous components of applied fertilizer should appear as linear, squared, and interacted terms. The model, with just the nitrogen and phosphorous terms included, would be estimated as follows:

$$Y = \alpha_1 + \beta_1 N + \beta_2 N^2 + \beta_3 P + \beta_4 P^2 + \beta_5 N*P$$

However, in addition to the difficulties in estimating the response to applied phosphorous, where absorption and the current stock of nutrients in the soil are also important for response (see x), the collinearity between applied nitrogen and phosphorous create additional difficulties in estimating the complete quadratic production function. In an experimental context, the researcher can systematically vary the amount of applied nitrogen and phosphorous used by varying the type of fertilizer applied. When using household data, as I do here, the researcher must use the variation provided by households and their choices of fertilizer type(s).

In Kenya, households overwhelmingly choose to use one of two regimes of fertilization on maize: (1) DAP only or (2) DAP and CAN together in some relatively fixed proportion. These two schemes result in a high degree of collinearity between applied nitrogen and phosphorous. The graphs in Figure A.1 show how applied nitrogen and phosphorous vary at the field level by zone. Notice the several prominent lines in each of the graphs. What is misleading about these graphs, however, is the fact that they do not show the density of observations at an individual point and along an individual line. There are a large number of observations on each of the “lines” found in these plots. I show this by creating an additional set of histograms (Figure A.2) showing the ratio of applied phosphorous to nitrogen at the field level. The large number of observations at 1.1 represents fields with DAP only. The pile of observations around 0.5 represents fields fertilized with DAP and CAN in relatively fixed proportions.

Figure A.1 Scatter plots of applied nitrogen and applied phosphorous

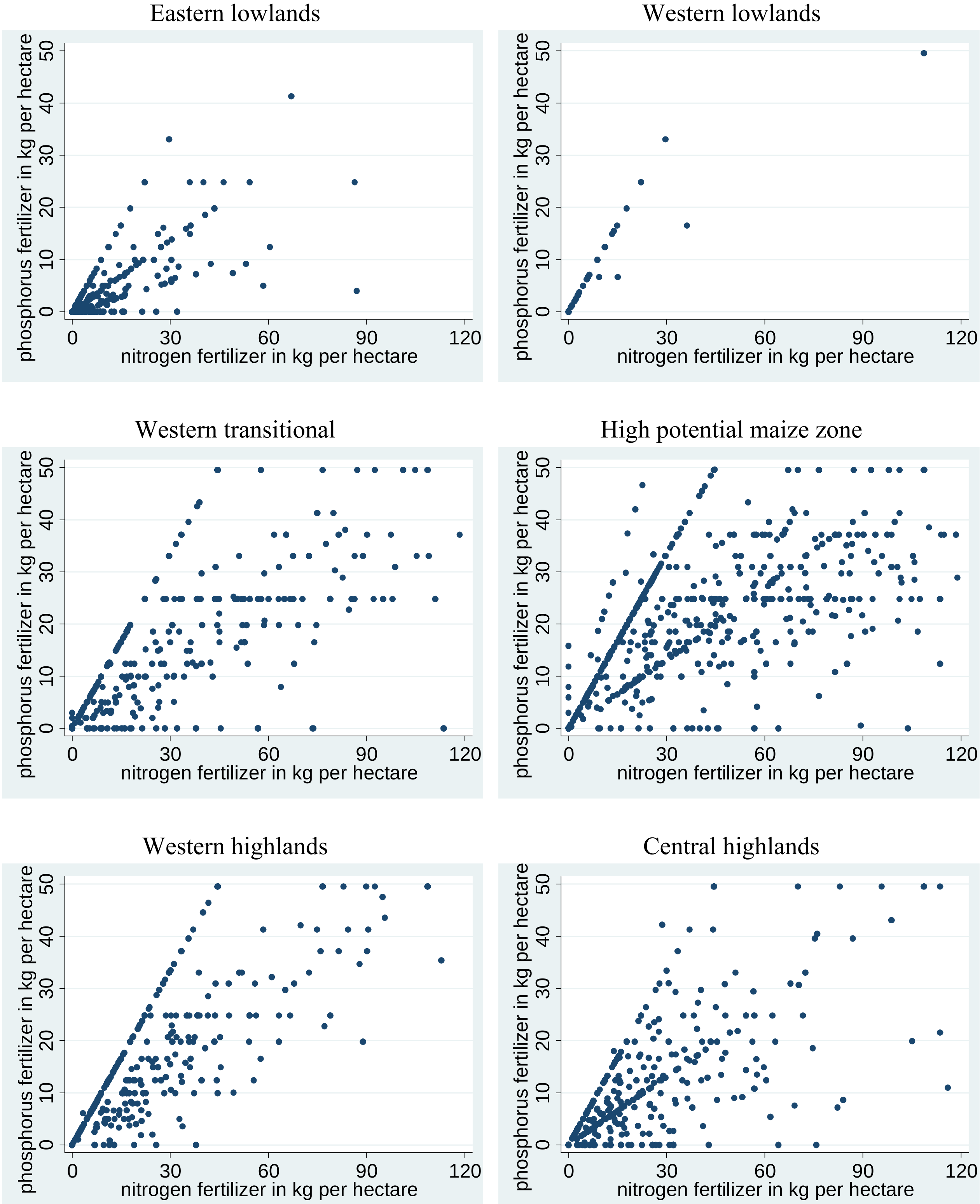
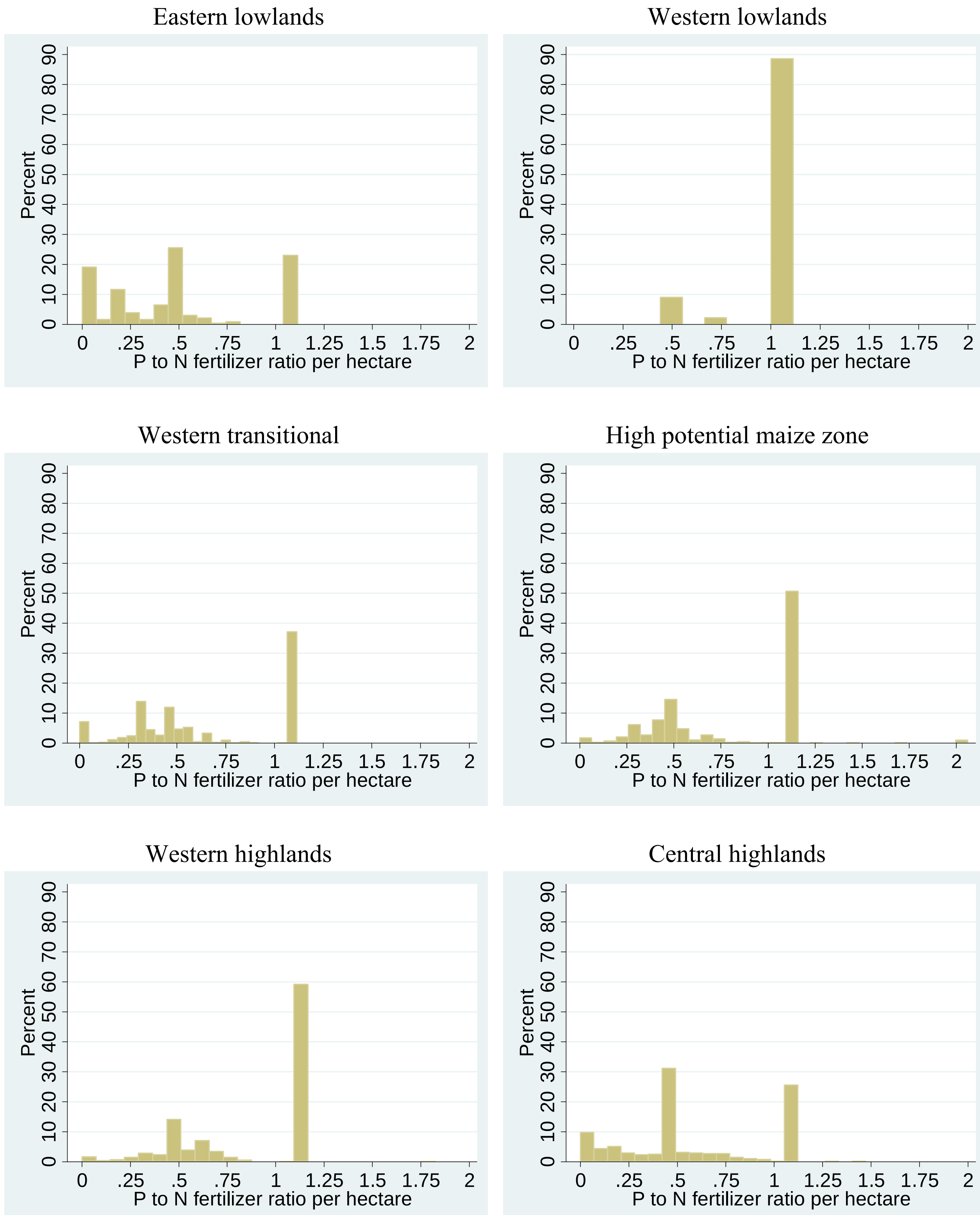


Figure A.2 Histograms of applied phosphorous to applied nitrogen



Given these trends, estimating the complete quadratic model is problematic. Insufficient variation between the nitrogen and phosphorous variables biases the estimates in the complete quadratic model and makes recovery of response to either nutrient very difficult. When including the two nutrients together in the same model, the coefficients are highly sensitive to functional form and sample selection. As such, I drop from my model the linear and squared phosphorous terms, but keep the interaction between nitrogen and phosphorous:

$$Y = \alpha_1 + \beta_1 N + \beta_2 N^2 + \beta_3 N*P$$

The applied nitrogen variable then acts as proxy for overall fertilizer application. The fact that my estimates of marginal product of nitrogen are similar to others found in the literature, both from household data and experiment station trials, provides added validity to my fertilizer modeling strategy.

Appendix 4: Detail on process for grouping soils for nitrogen interactions

In this appendix, I detail how I arrived at the soil groups used in this analysis and the alternatives that I forewent along the way. For this study, data on time invariant soil characteristics (i.e., drainage, depth, texture) and FAO soil classifications are available at the village level from the Kenya Soil Survey and the Ministry of Agriculture from data originally collected in 1980 (see Figure A.4 in Appendix 12 for a map). While the observed soil characteristics do not include soil fertility or soil organic matter levels per se, some of what is observed is likely correlated with those important soil variables. For example, soils with more clay are more likely to have higher soil organic matter levels than sandier soils, and clay is more likely to hold onto applied fertilizer than sand (see Sileshi et al. 2010). Soil depth could be an indicator of potential root depth, meaning deeper soils could yield higher growth levels (see Feller and Bearer 1997). Soil drainage is necessary for processing organic matter.

My first attempt at grouping soils involved multivariate cluster analysis on the observable soil characteristics at the village level: soil drainage, soil depth, clay, silt, and sand content. Multivariate cluster analysis uses the natural grouping of villages based on similarities in the given characteristics by minimizing the Euclidean distance within those variables. I utilized group averaged hierarchical cluster analysis given well-noted problems with other hierarchical methods (Cunningham and Ogilvie 1972; Milligan 1980) and chose to partition the data into the "optimal" number of groups using the Duda and Hart index (Duda and Hart 1973). No matter the number of groups I choose, however, the groupings did not seem to adequately capture the variation in soil characteristics that contributed to observed levels of fertilizer use and yield.

Instead, I moved to grouping soils manually, relying on information revealed through cluster analysis. The first manual set of groups I created was done by focusing exclusively on the clay, sand, and silt composition of the soils given these variables are most related to soil organic matter. In this data set, variation in soil composition is not immense, with over 50 percent of villages having exactly 70 percent clay content and about 80 percent of villages having at least 50 percent clay. Not only that, but advice from soil scientists suggested that within village variation in soil composition can be considerable, so using village level data on soil composition for pooling likely would generate significant measurement error. Because soil depth and drainage variables are not nearly as important for fertilizer response and are not as correlated with soil organic matter, using these two variables to motivate soil groups is not appropriate.

With the problems associated with grouping variables using the time invariant soil characteristics at the village level, I turned to the FAO soil types instead. Observed, again, at the village level, the FAO soil types represent an attempt to classify soils based on their soil formation process “defined in terms of diagnostic horizons, properties and materials, which to the greatest extent possible should be measurable and observable in the field” (IUSS Working Group WRB 2007). I then moved to grouping soils based on FAO soil type using (1) key terms in their definitions then (2) the over-arching groups detailed in the IUSS Working Group Report (see Table 1). Data on the landform in each village also aided categorization. For example, a large number of villages are found on volcanic footridges and plains, meaning the volcanic attributes of the soils lead to the creation of the first soil group. Given the need for sufficient variation within and between groups, care was taken to include a large number of households (no less than 100 households) in each soil group. Furthermore, some soil types were moved between groups for testing (i.e., do Cambisols belong with the soils found in volcanic areas or the high humus areas?) in order to arrive at the final groups in Table 14. While more data on soil type and quality would improve estimation, this grouping scheme represents a best effort to use the available data and consider knowledge of soil properties and fertilizer response.

Appendix 5: Descriptive statistics of variables included in the production function

Table A.2 Distribution of variables in the production function

Variable	Unit of Measure	Percentiles								
		1st	5th	10th	25th	50th	75th	90th	95th	99th
Yield per hectare	Kilograms of maize per hectare (using Liu and Myers output index)	138	444	716	1342	2375	3686	5150	6140	8099
Nitrogen fertilizer per hectare	Kilograms of nitrogen from inorganic fertilizer applied per hectare	0	0	0	2	20	38	65	79	109
Phosphorous fertilizer per hectare	Kilograms of phosphorus from inorganic fertilizer applied per hectare	0	0	0	0	12	25	33	40	50
Seed rate per hectare	Kilograms of maize seed used per hectare	5	10	11	20	25	25	31	37	49
Hectares	Total hectares on maize field	0.1	0.1	0.1	0.2	0.4	0.8	1.2	2.0	3.2
Manure or compost	Binary variable	1= manure or compost (30%) 0=none (70%)								
Hybrid seed	Binary variable	1= new hybrid (76%) 0= other seed type (24%)								
Intercropped with legume	Binary variable	1= yes (14%) 0= no (86%)								
Crops per field	Number of crops (including maize) on field	1=(14%) 2=(45%) 3=(15%) 4=(9%) 5=(7%) 6=(5%) 7=(4%)								
Rain stress	Proportion of 20-day periods when rainfall was less than 40 mm during the main growing season	0	0	0	0.1	0.2	0.4	0.6	0.7	0.8
Asset wealth	Real KSH value of representative group of assets (in 1000 KSH)	9	48	74	142	272	511	980	1407	2793

Note: Includes all 906 households and 4717 fields in production function estimation. Refer to Table 5 for variable units and Chapter 4 for more about what these variables measure.

Table A.3 Standard deviation of variables in the production function split by zone group

	Total sample n=906			Households in lowlands areas n=144			Households in high potential and transitional areas n=495			Households in highlands areas n=267		
	overall	between	within	overall	between	within	overall	between	within	overall	between	within
Yield	1778	1076	1422	1499	826	1261	1760	1032	1434	1769	994	1489
Nitrogen	26.5	29.6	17.5	11.9	7.5	9.6	28.2	20.6	19.3	22.6	15.0	17.2
Phosphorous	13.5	10.2	8.8	6.1	3.7	5.0	13.3	9.6	9.2	12.7	8.3	9.6
Seed rate	8.4	4.9	7.0	10.3	6.0	8.5	7.0	3.5	6.0	9.3	5.5	7.8
Hectares	0.65	0.50	0.42	0.46	0.31	0.33	0.76	0.57	0.51	0.23	0.14	0.19
Manure	0.46	0.31	0.34	0.49	0.31	0.38	0.39	0.23	0.33	0.48	0.32	0.36
Hybrid	0.43	0.33	0.28	0.45	0.24	0.38	0.33	0.24	0.24	0.40	0.31	0.28
Legume	0.35	0.21	0.28	0.47	0.31	0.37	0.31	0.17	0.27	0.30	0.17	0.25
Crops per field	1.6	0.8	1.4	1.8	0.8	1.6	1.5	0.7	1.3	1.5	0.9	1.3
Rainfall stress	0.22	0.18	0.13	0.22	0.16	0.16	0.20	0.15	0.13	0.23	0.21	0.12
Asset wealth	517	543	0	390	352	0	496	520	0	607	664	0

Note: "n" refers to the number of households in each group by number of years using fertilizer on any maize field; however, standard deviations are computed at the field level. Zone groups are defined in Section 2 of Chapter 4. "Overall" refers to standard deviation over entire sample in group. "Within" refers to standard deviation from the household level mean (as calculated for Mundlak-Chamberlain). "Between" refers to the standard deviation across households in a given year then averaged across all three survey years. There is no "within" variation in asset wealth because the variable is measuring a household average over time. Refer to Table 5 for variable units.

Table A.4 Averages of select production function variables by district and soil group

Province	District	Soil group	Yield (kg/ha)	N (kg/ha)	P (kg/ha)	P/N ratio	Fert fields (%)	Manure (%)	Hybrid (%)	Rain total (mm)	Rain stress (%)
Coast	Kilifi	3	1336	7.4	5.3	0.60	10	29	32	252	56
	Kwale	6	1156	0.9	0.4	0.44	2	29	9	242	69
	Taita Tav.	5	949	-	-	-	0	31	26	283	50
Eastern	Kitui	3	1312	-	-	-	0	34	12	289	51
	Machakos	3	1900	12.9	7.7	0.76	43	59	16	313	47
	Makueni	3	1607	14.9	5.1	0.36	62	70	46	271	49
	Meru	1	3145	25.2	14.8	0.66	89	60	98	545	27
	Mwingi	2	1703	16.6	10.4	0.60	10	69	30	326	40
	Mwingi	3	2229	10.8	10.6	0.90	19	68	45	334	38
Nyanza	Kisii	2	2242	29.2	22.5	0.93	98	10	93	889	12
	Kisii	4	2309	24.2	19.1	0.92	98	6	89	858	14
	Kisumu	5	1204	15.3	10.8	0.75	3	16	32	719	12
	Siaya	3	1574	13.9	11.8	1.0	14	36	8	710	16
	Siaya	4	2008	19.4	16.3	1.0	24	49	23	719	16
	Siaya	5	1431	5.8	3.7	0.59	3	12	10	655	19
Western	Bungoma	2	2724	37.3	23.8	0.80	91	18	90	848	6
	Bungoma	3	3507	45.7	24.2	0.68	90	17	96	828	8
	Bungoma	4	2733	45.6	22.3	0.68	89	18	94	805	6
	Kakamega	2	3864	64.4	29.3	0.54	96	15	91	746	11
	Kakamega	3	2508	45.7	22.4	0.64	64	25	92	876	4
	Kakamega	4	2453	24.2	10.5	0.56	54	46	44	869	5
	Vihiga	3	2689	26.5	13.9	0.67	71	33	48	891	7
	Vihiga	4	2795	26.1	14.8	0.71	87	39	57	893	8
Central	Muranga	1	2554	28.0	13.4	0.59	91	55	69	378	60
	Muranga	4	2598	19.1	15.4	0.83	87	50	63	377	56
	Nyeri	1	3110	31.6	11.0	0.44	93	68	78	381	54
	Nyeri	2	2807	33.6	12.6	0.41	68	59	68	348	58
Rift Valley	Bomet	1	3119	21.7	23.3	1.1	100	9	97	858	22
	Nakuru	1	2891	23.6	22.0	1.0	94	18	98	538	40
	Nakuru	2	1775	20.1	17.3	1.0	72	17	52	497	50
	Nakuru	3	3012	20.5	20.6	1.1	97	18	92	527	36
	Narok	1	3029	11.6	12.3	1.1	28	10	99	469	56
	Narok	2	3277	11.1	12.4	1.1	3	9	99	484	55
	Trans Nz.	4	3805	53.8	27.1	0.61	89	17	94	676	18
	Uasin Gis.	1	3585	36.5	20.5	0.63	86	10	95	618	24
	Uasin Gis.	2	3048	51.4	25.6	0.62	95	14	91	600	28
	Laikipia	2	2125	15.0	13.7	0.98	4	56	66	285	62
	Laikipia	5	2207	-	-	-	0	45	48	289	60

Note: N, P, and P to N ratio values represent averages across fertilizer users (excludes non-users). District and soil group combinations in gray are excluded from estimation due to (1) very low rainfall, (2) poor soil conditions (i.e., soil groups 5 and 6) or (3) practically no fertilizer users (i.e., less than 10 percent fertilized fields). For information on soil groups, see Table 14.

Appendix 6: Modified quadratic production function results

Table A.5 Production function regression results

	Pooled OLS	FE	CRE
N*zone1	34.52**	14.65	25.45
	(17.19)	(16.33)	(17.46)
N*zone2	22.73***	16.20***	17.58***
	(4.794)	(5.072)	(4.901)
N*zone3	18.22***	9.346	14.10**
	(6.621)	(7.842)	(6.631)
N*N*zone1	-0.781***	-0.694***	-0.724***
	(0.205)	(0.184)	(0.210)
N*N*zone2	-0.122***	-0.0585	-0.0938**
	(0.0441)	(0.0499)	(0.0463)
N*N*zone3	-0.0926	-0.112	-0.0889
	(0.0799)	(0.0927)	(0.0812)
N*P*zone1	1.391***	1.328***	1.379***
	(0.394)	(0.372)	(0.417)
N*P*zone2	0.291***	0.230***	0.256***
	(0.0797)	(0.0771)	(0.0780)
N*P*zone3	0.193	0.392**	0.218
	(0.142)	(0.166)	(0.148)
N*soil1	-2.944	-4.266	-2.712
	(3.836)	(4.491)	(3.825)
N*soil2	2.871	-2.524	2.317
	(3.158)	(3.603)	(3.106)
N*soil3	-4.715	-4.201	-4.733
	(3.164)	(3.482)	(3.165)
N*soil4	omitted	omitted	omitted
N*rain*zone1	34.72*	58.86***	41.00*
	(20.99)	(20.58)	(21.43)
N*rain*zone2	-20.64***	-16.42**	-18.66**
	(7.340)	(8.214)	(7.419)
N*rain*zone3	17.01*	25.74**	17.82*
	(9.190)	(11.30)	(9.219)
seed	61.76***	55.75***	57.17***
	(9.725)	(10.59)	(9.803)
seed*seed	-0.524***	-0.470**	-0.495**
	(0.196)	(0.206)	(0.194)
hect	-751.2***	-961.0***	-944.5***
	(96.50)	(110.2)	(99.94)
hect*hect	126.7***	141.4***	135.9***
	(21.82)	(24.85)	(21.29)

Table A.5 (cont'd)

asset	0.628***		0.526***
	(0.137)		(0.140)
asset*asset	-0.000126**		-0.000105**
	(5.04e-05)		(5.12e-05)
rain	-1,258***	-1,431***	-1,457***
	(270.0)	(271.3)	(269.9)
manure	212.5***	179.0***	189.2***
	(58.32)	(63.44)	(64.16)
hybrid	579.7***	633.6***	568.7***
	(86.57)	(100.5)	(95.25)
hybrid*rain	-322.2	-553.1**	-307.5
	(250.1)	(273.3)	(250.9)
legume	-211.6***	-93.39	-97.99
	(80.94)	(79.31)	(79.23)
crop1 (monocropped)	omitted	omitted	omitted
crop2	266.6***	424.6***	315.2***
	(64.11)	(66.66)	(63.36)
crop3	570.6***	697.5***	636.2***
	(81.27)	(89.96)	(81.44)
crop4	940.2***	1,180***	1,025***
	(101.9)	(103.6)	(101.4)
crop5	1,041***	1,210***	1,122***
	(104.8)	(110.0)	(105.6)
crop6	1,476***	1,543***	1,573***
	(131.7)	(130.6)	(131.8)
crop7	1,571***	1,774***	1,700***
	(140.2)	(140.0)	(142.6)
District dummy variables:			
Machakos	omitted		omitted
Makueni	-635.6**		-793.5***
	(248.6)		(279.0)
Meru	-553.2		-456.6
	(362.3)		(426.2)
Mwingi	-9.718		138.7
	(268.0)		(280.6)
Kisii	-722.7**		-591.2
	(285.3)		(402.0)
Siaya	-614.6**		-382.9
	(269.3)		(359.6)
Bungoma	-303.7		-80.43
	(310.4)		(436.7)
Kakamega	-349.8		-128.1
	(280.7)		(417.8)

Table A.5 (cont'd)

Vihiga	-446.4		-127.6
	(291.5)		(411.4)
Muranga	-694.7*		-980.6**
	(365.6)		(399.0)
Nyeri	-524.3		-762.1**
	(349.7)		(382.2)
Bomet	-757.0**		-655.5
	(374.0)		(449.6)
Nakuru	257.2		178.9
	(286.9)		(314.9)
Narok	771.7**		446.9
	(370.2)		(398.7)
Trans Nzoia	424.9		380.6
	(320.3)		(415.0)
Uasin Gishu	118.4		80.42
	(329.5)		(405.0)
FAO soil classification dummy variables:			
Cambisols	omitted		omitted
Phaeozems	303.9		267.1
	(193.9)		(187.6)
Luvisols	200.8		85.62
	(184.4)		(189.5)
Greyzems	1,311***		1,361***
	(421.1)		(414.3)
Podzols	516.9**		469.0*
	(240.7)		(246.1)
Regosols	1,104***		1,105***
	(230.6)		(224.7)
Rankers	418.3**		461.5***
	(173.5)		(170.5)
Year dummy variables:			
1997	omitted	omitted	omitted
2000	60.89	65.71	54.71
	(70.25)	(70.33)	(70.32)
2004	237.1***	209.7***	220.5***
	(59.29)	(60.72)	(59.07)
2007	719.8***	776.8***	733.1***
	(71.00)	(74.18)	(72.20)
2010	302.3***	357.1***	314.9***
	(91.04)	(95.19)	(92.09)
Mundlak-Chamberlain device:			
mean N			6.537*
			(3.514)

Table A.5 (cont'd)

mean P			1.859
			(7.232)
mean seed			6.801
			(7.457)
mean hect			279.3***
			(95.16)
mean rain			1,240*
			(653.1)
mean manure			69.76
			(135.6)
mean hybrid			-35.85
			(153.4)
mean legume			-411.7**
			(197.5)
mean crop			-73.55
			(45.52)
constant	96.42	798.2***	-199.4
	(363.6)	(181.8)	(551.7)
Number of fields	4714	4714	4714
Number of households	906	906	906
R-squared	0.350	0.280	0.358

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.6 Marginal effects of the production function

	Pooled OLS	CRE
Nitrogen per hectare	20.69***	16.65***
	(2.143)	(2.220)
Phosphorous per hectare	7.646***	7.183***
	(1.734)	(1.778)
Rain stress	-1,732***	-1,873***
	(225.7)	(227.5)
Seed per hectares	38.21***	34.92***
	(3.013)	(3.275)
Hectares in field	-596.5***	-778.6***
	(73.46)	(79.42)
Asset wealth per hectare	0.518***	0.435***
	(0.100)	(0.103)
Manure	212.5***	189.2***
	(58.32)	(64.16)
Hybrid	502.3***	494.9***
	(65.18)	(73.59)
Legume	-211.6***	-97.99
	(80.94)	(79.23)
Two crops on field	266.6***	315.2***
	(64.11)	(63.36)
Three crops on field	570.6***	636.2***
	(81.27)	(81.44)
Four crops on field	940.2***	1,025***
	(101.9)	(101.4)
Five crops on field	1,041***	1,122***
	(104.8)	(105.6)
Six crops on field	1,476***	1,573***
	(131.7)	(131.8)
Seven crops on field	1,571***	1,700***
	(140.2)	(142.6)
2000	60.89	54.71
	(70.25)	(70.32)
2004	237.1***	220.5***
	(59.29)	(59.07)
2007	719.8***	733.1***
	(71.00)	(72.20)
2010	302.3***	314.9***
	(91.04)	(92.09)

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

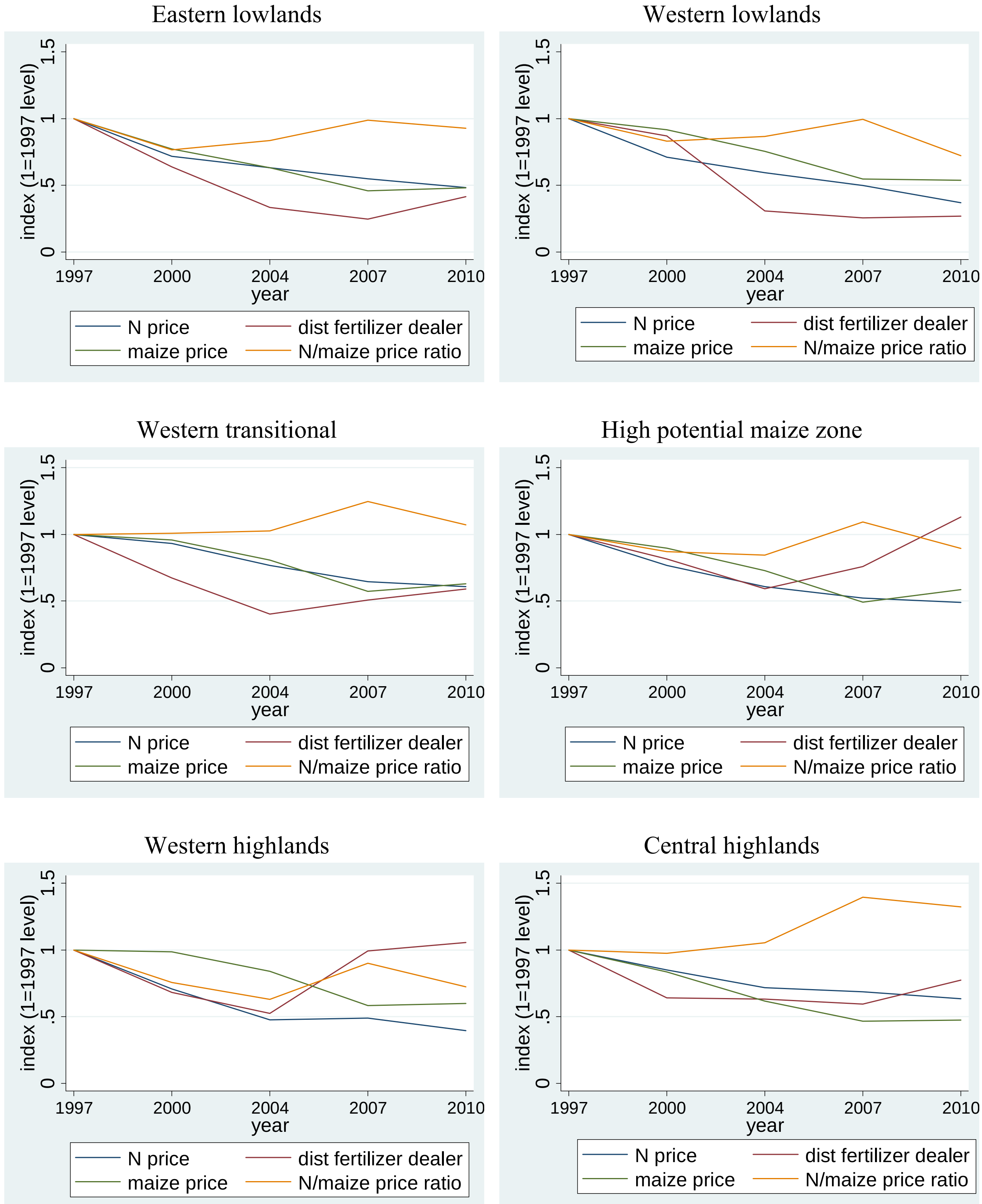
Appendix 7: Variables related to fertilizer profitability scenarios

Table A.7 Averages of variables related to fertilizer profitability by district and soil group

Province	District	Soil group	Dist fert (km)	N market (KSH)	Maize sell (KSH)	Maize buy (KSH)	Net sell (%)	Scenario 1		Scenario 5	
								MVCR	AVCR	MVCR	AVCR
Coast	Kilifi	3	4.0	233	25.8	32.7	0	-	-	-	-
	Kwale	6	25.9	249	27.5	42.3	2	-	-	-	-
	Taita Tav.	5	12.8	-	23.5	31.5	0	-	-	-	-
Eastern	Kitui	3	4.1	213	22.1	31.2	11	-	-	-	-
	Machakos	3	3.9	241	21.9	31.1	0	3.7	4.2	3.8	4.3
	Makueni	3	2.3	163	23.2	28.2	3	5.1	5.9	4.5	5.2
	Meru	1	1.3	216	22.9	28.7	10	1.9	2.1	1.8	2.0
	Mwingi	2	1.8	225	23.5	32.1	0	5.2	5.6	5.4	5.8
	Mwingi	3	2.0	232	23.7	32.3	0	4.4	5.2	4.6	5.4
Nyanza	Kisii	2	1.2	219	25.7	28.0	3	2.1	2.4	1.9	2.1
	Kisii	4	1.5	220	25.4	27.7	9	1.9	2.1	1.7	1.9
	Kisumu	5	3.7	287	25.5	26.4	0	-	-	-	-
	Siaya	3	5.2	347	24.4	29.4	4	2.1	2.6	1.9	2.4
	Siaya	4	3.1	356	24.3	29.3	10	2.3	3.0	1.9	2.5
	Siaya	5	4.7	348	24.2	28.9	0	-	-	-	-
Western	Bungoma	2	1.7	242	22.7	28.1	12	1.7	2.0	1.7	2.1
	Bungoma	3	6.2	243	23.1	28.6	67	0.8	1.2	0.8	1.1
	Bungoma	4	1.8	244	23.3	28.9	9	1.4	1.7	1.4	1.7
	Kakamega	2	2.1	241	22.6	25.8	55	1.3	1.8	1.2	1.7
	Kakamega	3	5.7	241	22.7	26.0	8	1.0	1.3	0.9	1.1
	Kakamega	4	1.8	242	22.7	26.1	0	1.4	1.5	1.4	1.5
	Vihiga	3	2.9	266	24.9	27.6	0	0.9	1.1	0.8	0.9
	Vihiga	4	3.7	263	24.5	27.0	0	1.3	1.5	1.1	1.3
Central	Muranga	1	0.7	221	24.3	29.7	16	2.2	2.5	2.3	2.5
	Muranga	4	1.1	209	22.0	27.7	0	2.5	2.7	2.4	2.6
	Nyeri	1	1.2	213	25.4	29.4	3	2.3	2.6	2.0	2.3
	Nyeri	2	2.4	209	24.5	28.6	0	3.0	3.2	2.6	2.8
Rift Valley	Bomet	1	1.7	338	26.8	27.5	8	1.1	1.3	1.0	1.2
	Nakuru	1	4.2	272	21.1	26.7	16	0.7	0.9	0.7	0.9
	Nakuru	2	2.6	272	21.3	27.1	15	1.0	1.1	0.9	1.1
	Nakuru	4	3.0	271	21.1	26.7	18	1.0	1.1	0.9	1.1
	Narok	1	5.1	292	20.7	25.5	71	0.4	0.5	0.4	0.5
	Narok	2	4.6	303	21.8	27.6	21	-	-	-	-
	Trans Nz.	4	2.4	186	20.6	25.7	27	1.3	1.8	1.1	1.6
	Uasin Gis.	1	5.8	215	21.3	24.4	54	0.9	1.2	0.8	1.1
	Uasin Gis.	2	3.5	215	21.1	24.5	4	1.2	1.6	1.1	1.5
	Laikipia	2	3.6	183	20.8	26.6	11	-	-	-	-
	Laikipia	5	1.9	186	22.9	28.8	4	-	-	-	-

Note: Values averaged over 2004, 2007, 2010. Net seller represents the percent of households that are consistently net sellers across all surveys. Gray areas excluded from analysis.

Figure A.3 Plots of changes in relative accessibility of fertilizer over survey years



Note: All values indexed to observed levels in 1997. All prices adjusted to 2010 levels using the CPI.

Appendix 8: Optimal and actual nitrogen use rates

Table A.8 Estimated optimal versus actual nitrogen use rates by district and soil group

Province	District	Soil group	Nitrogen application rates (kilograms/hectare) mean (standard deviation)						
			Estimated optimal levels		Actual observed levels				
			MVCR=2	MVCR=1	1997	2000	2004	2007	2010
Eastern	Machakos	3	24.7 (8.7)	32.3 (9.0)	3.9 (3.4)	3.2 (2.0)	13.4 (16.9)	11.4 (13.3)	21.1 (22.3)
	Makueni	3	25.9 (5.6)	31.6 (5.8)	8.4 (7.8)	13.7 (15.7)	10.5 (10.4)	15.6 (16.2)	25.1 (14.3)
	Meru	1	17.9 (20.4)	70.7 (18.0)	24.7 (21.7)	24.3 (18.4)	24.9 (20.2)	27.6 (19.6)	29.7 (18.5)
	Mwingi	2	37.8 (8.7)	44.0 (8.9)	2.3 (0.2)	5.4 (1.4)	22.2 (0)	13.3 (12.6)	29.5 (22.6)
	Mwingi	3	27.1 (13.9)	33.6 (13.5)	1.8 (0)	11.1 (0)	3.2 (1.2)	13.1 (12.9)	22.2 (7.4)
Nyanza	Kisii	2	23.1 (21.3)	76.1 (18.3)	20.8 (14.9)	16.9 (9.9)	36.7 (36.3)	27.5 (13.6)	39.3 (25.1)
	Kisii	4	12.8 (16.2)	62.9 (15.9)	14.6 (10.2)	15.7 (10.9)	23.2 (17.1)	26.5 (16.5)	40.8 (24.5)
	Siaya	3	10.7 (11.2)	21.3 (10.3)	0	0	8.6 (4.4)	6.5 (3.1)	19.7 (28.9)
	Siaya	4	14.6 (11.9)	26.6 (11.4)	0.7 (0)	15.3 (20.3)	11.1 (7.7)	11.9 (7.6)	36.3 (42.3)
Western	Bungoma	2	22.2 (20.0)	76.9 (16.8)	22.4 (11.6)	33.3 (20.3)	34.0 (22.6)	51.4 (29.7)	42.7 (27.2)
	Bungoma	3	0.1 (0.4)	26.6 (14.7)	38.1 (27.8)	38.5 (24.5)	57.0 (19.2)	41.2 (23.3)	43.4 (22.5)
	Bungoma	4	12.8 (15.4)	63.1 (15.9)	32.1 (25.7)	34.8 (24.3)	48.1 (29.3)	53.8 (29.5)	56.1 (26.5)
	Kakamega	2	11.9 (11.6)	70.6 (14.1)	46.9 (21.1)	64.2 (23.5)	72.3 (28.1)	55.5 (23.8)	66.7 (21.1)
	Kakamega	3	0.2 (1.7)	32.3 (16.7)	31.9 (24.0)	30.8 (26.1)	49.2 (32.9)	52.4 (20.8)	51.1 (27.4)
	Kakamega	4	1.0 (3.5)	38.9 (13.7)	45.6 (30.0)	18.3 (20.7)	27.3 (23.7)	25.0 (22.5)	21.7 (17.2)
	Vihiga	3	0.1 (0.5)	9.9 (13.0)	11.2 (9.9)	18.4 (20.4)	28.3 (24.7)	28.4 (23.0)	34.3 (29.4)
	Vihiga	4	1.7 (5.3)	30.5 (22.0)	16.5 (21.2)	26.4 (18.1)	25.1 (21.4)	24.2 (22.9)	34.8 (25.44)

(continued on next page)

Table A.8 (cont'd)

Central	Muranga	1	33.6 (16.1)	84.6 (13.1)	38.0 (35.2)	31.6 (27.1)	22.1 (18.9)	15.3 (10.9)	37.7 (25.0)
	Muranga	4	31.9 (7.4)	90.1 (3.9)	18.4 (18.1)	23.9 (8.7)	12.3 (10.5)	17.8 (6.0)	17.2 (9.8)
	Nyeri	1	28.4 (16.4)	81.9 (14.4)	29.9 (21.6)	30.8 (26.7)	37.3 (25.9)	26.4 (18.7)	30.1 (24.9)
	Nyeri	2	45.9 (19.9)	105.7 (14.8)	34.8 (31.1)	27.5 (17.2)	25.0 (16.2)	27.3 (22.6)	34.9 (34.9)
Rift Valley	Bomet	1	0.4 (2.8)	22.1 (16.4)	26.1 (11.0)	19.5 (7.8)	20.8 (9.2)	18.7 (9.4)	22.1 (9.4)
	Nakuru	1	0	5.8 (7.4)	22.0 (7.3)	22.7 (11.9)	23.6 (16.4)	22.8 (13.1)	34.5 (18.4)
	Nakuru	2	0.1 (0.2)	16.9 (17.2)	19.7 (19.7)	17.3 (11.0)	22.8 (17.9)	22.7 (12.2)	18.5 (5.2)
	Nakuru	4	0.2 (2.0)	13.7 (15.2)	20.5 (9.2)	19.9 (7.5)	21.6 (12.5)	17.3 (6.8)	25.4 (16.6)
	Narok	1	0	0	11.1 (0)	11.5 (8.9)	13.1 (6.2)	9.3 (6.3)	15.9 (8.3)
	Trans Nz.	4	7.5 (10.2)	57.0 (14.0)	40.0 (22.1)	53.8 (26.4)	55.1 (22.3)	59.6 (26.9)	52.9 (23.3)
	Uasin Gis.	1	0.5 (3.8)	22.7 (15.7)	23.2 (12.4)	32.8 (15.5)	36.4 (21.2)	47.4 (23.2)	40.1 (26.3)
	Uasin Gis.	2	7.2 (12.6)	54.9 (18.2)	29.8 (15.1)	49.8 (25.4)	51.1 (25.9)	64.4 (29.8)	55.7 (28.7)

Note: The "estimated optimal" columns show the mean and standard deviation at the district and soil group level as computed using production function estimates and the relative acquisition price of nitrogen (market price plus transport cost) to the price of maize specific to the household (depending on net buyer or seller behavior). I compute this value to satisfy both MVCR=2 and MVCR=1 to compare the two levels. For instances where the value is zero, this means that no positive value of nitrogen application satisfies the requirement for that MVCR level. Negative optimal use values at the field level are replaced with zeros before averaging. "Average observed" values only include observations where fertilizer was applied. See text for more.

Table A.9 Nitrogen profitability and current use levels by district and soil group

Region	District	Soil group	Mean across survey years				Estimated optimal N (kg/ha)		Mean observed N (kg/ha) (excludes zeros)					% maize fields with fert				
			MP	AP	MVCR	AVCR	MVCR=2	MVCR=1	1997	2000	2004	2007	2010	1997	2000	2004	2007	2010
Eastern	Machakos	3	41	44	3.5	4.2	24.7	32.3	3.9	3.2	13.4	11.4	21.1	24	17	58	67	80
	Makueni	3	36	42	4.3	5.2	25.9	31.6	8.4	13.7	10.5	15.6	25.1	39	36	77	70	81
	Meru	1	18	20	1.8	2.1	17.9	70.7	24.7	24.3	24.9	27.6	29.7	89	93	95	90	89
	Mwingi	2	48	55	5.4	6.5	37.8	44.0	2.3	5.4	22.2	13.3	29.5	14	9	4	11	19
	Mwingi	3	42	50	4.7	5.6	27.1	33.6	1.8	11.1	3.2	13.1	22.2	11	7	29	14	30
Nyanza	Kisii	2	18	21	1.8	2.1	23.1	76.1	20.8	16.9	36.7	27.5	39.3	86	100	100	100	97
	Kisii	4	16	18	1.7	1.9	12.8	62.9	14.6	15.7	23.2	26.5	40.8	89	98	99	100	97
	Siaya	3	29	36	1.9	2.4	10.7	21.3	0	0	8.6	6.5	19.7	0	0	9	28	33
	Siaya	4	32	41	1.9	2.5	14.6	26.6	0.7	15.3	11.1	11.9	36.3	7	14	20	47	38
Western	Bungoma	2	18	21	1.7	2.1	22.2	76.9	22.4	33.3	34.0	51.4	42.7	86	88	96	95	93
	Bungoma	3	9	13	0.77	1.1	0.1	26.6	38.1	38.5	57.0	41.2	43.4	79	100	79	100	100
	Bungoma	4	14	18	1.3	1.7	12.8	63.1	32.1	34.8	48.1	53.8	56.1	73	88	96	93	93
	Kakamega	2	14	19	1.1	1.6	11.9	70.6	46.9	64.2	72.3	55.5	66.7	88	96	97	93	100
	Kakamega	3	10	14	0.8	1.1	0.2	32.3	31.9	30.8	49.2	52.4	51.1	32	57	67	78	81
	Kakamega	4	15	16	1.3	1.5	1.0	38.9	45.6	18.3	27.3	25.0	21.7	19	62	58	75	63
	Vihiga	3	9	11	0.7	0.9	0.1	9.9	11.2	18.4	28.3	28.4	34.3	53	52	71	87	86
	Vihiga	4	14	16	1.1	1.3	1.7	30.5	16.5	26.4	25.1	24.2	34.8	53	71	100	93	94
Central	Muranga	1	20	23	2.2	2.4	33.6	84.6	38.0	31.6	22.1	15.3	37.7	95	96	89	93	81
	Muranga	4	24	26	2.4	2.5	31.9	90.1	18.4	23.9	12.3	17.8	17.2	100	100	100	75	50
	Nyeri	1	19	22	2.0	2.3	28.4	81.9	29.9	30.8	37.3	26.4	30.1	86	88	97	96	96
	Nyeri	2	26	27	2.5	2.7	45.9	105.7	34.8	27.5	25.0	27.3	34.9	67	30	73	63	53
Rift Valley	Bomet	1	15	17	1.0	1.2	0.4	22.1	26.1	19.5	20.8	18.7	22.1	100	100	100	100	100
	Nakuru	1	9	11	0.6	0.8	0	5.8	22.0	22.7	23.6	22.8	34.5	97	92	95	94	85
	Nakuru	2	12	15	0.9	1.0	0.1	16.9	19.7	17.3	22.8	22.7	18.5	68	79	81	67	50
	Nakuru	3	12	14	0.9	1.0	0.2	13.7	20.5	19.9	21.6	17.3	25.4	95	96	98	98	96
	Narok	1	6	8	0.3	0.5	0	0	11.1	11.5	13.1	9.3	15.9	8	40	24	53	18
	Trans Nz.	4	11	16	1.1	1.6	7.5	57.0	40.0	53.8	55.1	59.6	52.9	69	89	92	90	94
	Uasin Gis.	1	9	13	0.8	1.1	0.5	22.7	23.2	32.8	36.4	47.4	40.1	54	88	92	91	94
	Uasin Gis.	2	12	17	1.0	1.5	7.2	54.9	29.8	49.8	51.1	64.4	55.7	88	98	95	96	98

Appendix 9: Reasons given by households that did not use fertilizer on maize

Table A.10 Reasons for not using fertilizer from villages included in analysis (white rows in Table A.7)

	Coast		Eastern		Nyanza		Western		Central		Rift Valley	
	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010
Not profitable			1	1	1			1		1		
Low response rate					1				1			2
Not enough cash; no cash when needed			17	8	18	11	24	24		7	8	9
Too expensive			3	14		6	2	6		1		
Maize price too low				2					1	1		
Fertilizer not available			1									
No need to use			18	12	3			4	3		3	10
Excessive vegetation												
Lack of advice on use						5				14		
Scorching effect; spoils the soil					1	1						
Low rains						1						
Practices organic farming								2		1		

Table A.11 Reasons for not using fertilizer from villages not included in analysis (gray rows in Table A.7)

	Coast		Eastern		Nyanza		Western		Central		Rift Valley	
	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010
Not profitable		2				4					2	
Low response rate					1	6					7	1
Not enough cash; no cash when needed	33	25	5	6	34	40					6	13
Too expensive	1	18	1	5		10						3
Maize price too low				5		1						
Fertilizer not available		2				1						
No need to use	27	12	10		41	39					13	29
Excessive vegetation					1							
Lack of advice on use	4	4			1							
Scorching effect; spoils the soil					1						3	
Low rains												
Practices organic farming												

Appendix 10: Descriptive statistics of variables included in binary response models

Table A.12 Mean and standard deviation of variables in binary response models

	Non-fertilized Fields (794 fields, 345 households)	Fertilized Fields (2727 fields. 788 households)
Age of household head (years)	58.2 (14.0)	55.6 (13.2)
Education of household head (years)	5.8 (3.9)	7.2 (4.5)
Sex of household head (1=female)	0.26 (0.44)	0.17 (0.37)
Farm size (hectares)	1.4 (1.2)	1.8 (2.5)
Own land with deed (1=yes)	0.53 (0.50)	0.58 (0.49)
Use manure or compost on fields (1=yes)	0.59 (0.49)	0.29 (0.45)
Use hybrid maize seed (1=yes)	0.36 (0.48)	0.86 (0.35)
Asset wealth (in 1000 KSH)	380 (700)	431 (690)
Successfully received credit (1=yes)	0.39 (0.49)	0.49 (0.50)
Distance to fertilizer seller (km)	3.7 (2.5)	2.7 (2.2)
Distance to extension service (km)	5.8 (4.7)	4.7 (4.1)
Part of a cooperative or group (1=yes)	0.71 (0.46)	0.79 (0.41)
Own a phone (1=yes)	0.36 (0.48)	0.49 (0.50)
Relative market price of N to maize	10.3 (2.5)	10.0 (1.9)
Received a gov't fertilizer subsidy (1=yes)	0.02 (0.13)	0.04 (0.20)
Indirectly affected by PEV (1=yes)	0.11 (0.31)	0.15 (0.36)
Directly affected by PEV (1=yes)	0.01 (0.08)	0.02 (0.13)

Note: Sample includes only households where fertilizer use is profitable on average (AVCR>1 for last four survey years). Only final four survey years included (1997 excluded).

Appendix 11: Binary response model estimates

Table A.13 Binary response model regression results

	Probit	Logit
Age of hh head	0.000556	0.00102
	(0.00260)	(0.00467)
Education of hh head	0.0388***	0.0722***
	(0.00854)	(0.0155)
Female hh head (1=yes)	-0.135*	-0.235*
	(0.0755)	(0.134)
Farm size (ha)	0.0447**	0.0935**
	(0.0223)	(0.0428)
Ratio of nitrogen price to maize price	-0.0530***	-0.0889***
	(0.0173)	(0.0312)
Own land (1=yes)	-0.0339	-0.0655
	(0.0628)	(0.113)
Manure or compost on field (1=yes)	-0.624***	-1.104***
	(0.0617)	(0.111)
New hybrid seed on field (1=yes)	0.956***	1.654***
	(0.0711)	(0.126)
Household asset level (1000 KSH)	1.43e-05	3.88e-05
	(7.23e-05)	(0.000139)
Obtained credit (1=yes)	0.0813	0.121
	(0.0637)	(0.116)
Distance to fertilizer dealer (km)	-0.0608***	-0.101***
	(0.0146)	(0.0264)
Distance to extension services (km)	0.00138	0.00283
	(0.00711)	(0.0128)
Member of cooperative or group (1=yes)	0.315***	0.576***
	(0.0707)	(0.126)
Own a cell phone (1=yes)	0.219**	0.380**
	(0.105)	(0.189)
Received gov't fertilizer subsidy (1=yes)	-1.397***	-2.507***
	(0.141)	(0.254)
Directly affected by PEV (1=yes)	0.502*	0.838*
	(0.267)	(0.481)
Indirectly affected by PEV (1=yes)	0.0676	0.138
	(0.123)	(0.223)
Rainfall stress	0.626***	1.322***
	(0.156)	(0.293)
Soil group 1	omitted	omitted
Soil group 2	-0.0196	-0.0710
	(0.103)	(0.193)
Soil group 3	-0.0651	-0.129

Table A.13 (cont'd)

	(0.116)	(0.214)
Soil group 4	0.0646	0.104
	(0.101)	(0.189)
Zone group 1	omitted	omitted
Zone group 2	1.051***	1.843***
	(0.101)	(0.180)
Zone group 3	1.087***	1.939***
	(0.112)	(0.201)
2000	omitted	omitted
2004	0.432***	0.777***
	(0.0840)	(0.152)
2007	0.324**	0.597**
	(0.141)	(0.254)
2010	-0.0220	-0.0641
	(0.141)	(0.256)
Constant	-0.661**	-1.319**
	(0.290)	(0.526)
Observations	3500	3500

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

Table A.14 Partial effects of binary response models

	LPM	Probit	Logit
Age of hh head	0.000153	0.000105	0.000108
	(0.000520)	(0.000491)	(0.000491)
Education of hh head	0.00757***	0.00734***	0.00760***
	(0.00160)	(0.00161)	(0.00162)
Female hh head (1=yes)	-0.0302*	-0.0262*	-0.0254*
	(0.0157)	(0.0150)	(0.0148)
Farm size (ha)	0.00364	0.00844**	0.00984**
	(0.00263)	(0.00421)	(0.00449)
Ratio of nitrogen price to maize price	-0.0134***	-0.0100***	-0.00936***
	(0.00360)	(0.00326)	(0.00327)
Own land (1=yes)	-0.00502	-0.00639	-0.00687
	(0.0123)	(0.0118)	(0.0119)
Manure or compost on field (1=yes)	-0.133***	-0.130***	-0.128***
	(0.0129)	(0.0136)	(0.0135)
New hybrid seed on field (1=yes)	0.250***	0.233***	0.226***
	(0.0161)	(0.0200)	(0.0199)
Household asset level (1000 KSH)	1.84e-06	2.70e-06	4.08e-06
	(1.26e-05)	(1.37e-05)	(1.46e-05)
Obtained credit (1=yes)	0.00959	0.0154	0.0127
	(0.0122)	(0.0120)	(0.0122)
Distance to fertilizer dealer (km)	-0.0137***	-0.0115***	-0.0106***
	(0.00298)	(0.00275)	(0.00276)
Distance to extension services (km)	0.000239	0.000261	0.000298
	(0.00141)	(0.00134)	(0.00134)
Member of cooperative or group (1=yes)	0.0652***	0.0623***	0.0633***
	(0.0144)	(0.0146)	(0.0144)
Own a cell phone (1=yes)	0.0396**	0.0412**	0.0398**
	(0.0200)	(0.0196)	(0.0197)
Received gov't fertilizer subsidy (1=yes)	-0.366***	-0.354***	-0.362***
	(0.0327)	(0.0403)	(0.0420)
Directly affected by PEV (1=yes)	0.0662	0.0821**	0.0773**
	(0.0466)	(0.0372)	(0.0384)
Indirectly affected by PEV (1=yes)	0.0162	0.0126	0.0143
	(0.0257)	(0.0227)	(0.0228)
Rain stress	0.174***	0.118***	0.139***
	(0.0314)	(0.0293)	(0.0306)
Soil group 1	omitted	omitted	omitted
Soil group 2	-0.0158	-0.00376	-0.00758
	(0.0185)	(0.0196)	(0.0206)
Soil group 3	-0.0135	-0.0126	-0.0139
	(0.0232)	(0.0227)	(0.0231)
Soil group 4	-0.00330	0.0121	0.0108

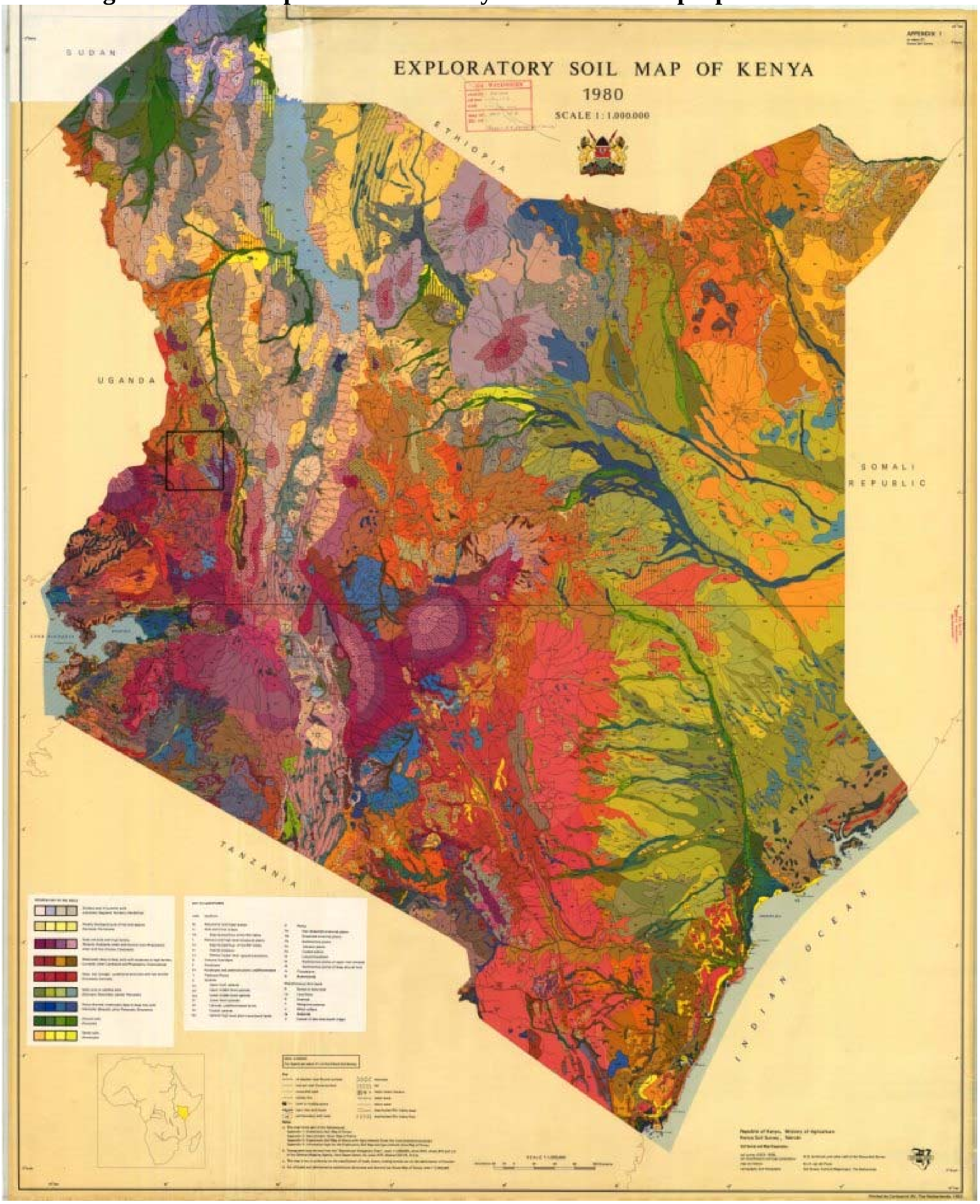
Table A.14 (cont'd)

	(0.0179)	(0.0189)	(0.0197)
Zone group 1	omitted	omitted	omitted
Zone group 2	0.337***	0.271***	0.270***
	(0.0224)	(0.0302)	(0.0307)
Zone group 3	0.328***	0.278***	0.280***
	(0.0240)	(0.0312)	(0.0316)
2000	omitted	omitted	omitted
2004	0.0891***	0.0808***	0.0808***
	(0.0164)	(0.0161)	(0.0162)
2007	0.0731***	0.0624**	0.0638**
	(0.0272)	(0.0269)	(0.0269)
2010	-3.90e-05	-0.00466	-0.00758
	(0.0288)	(0.0299)	(0.0302)
Observations	3500	3500	3500

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

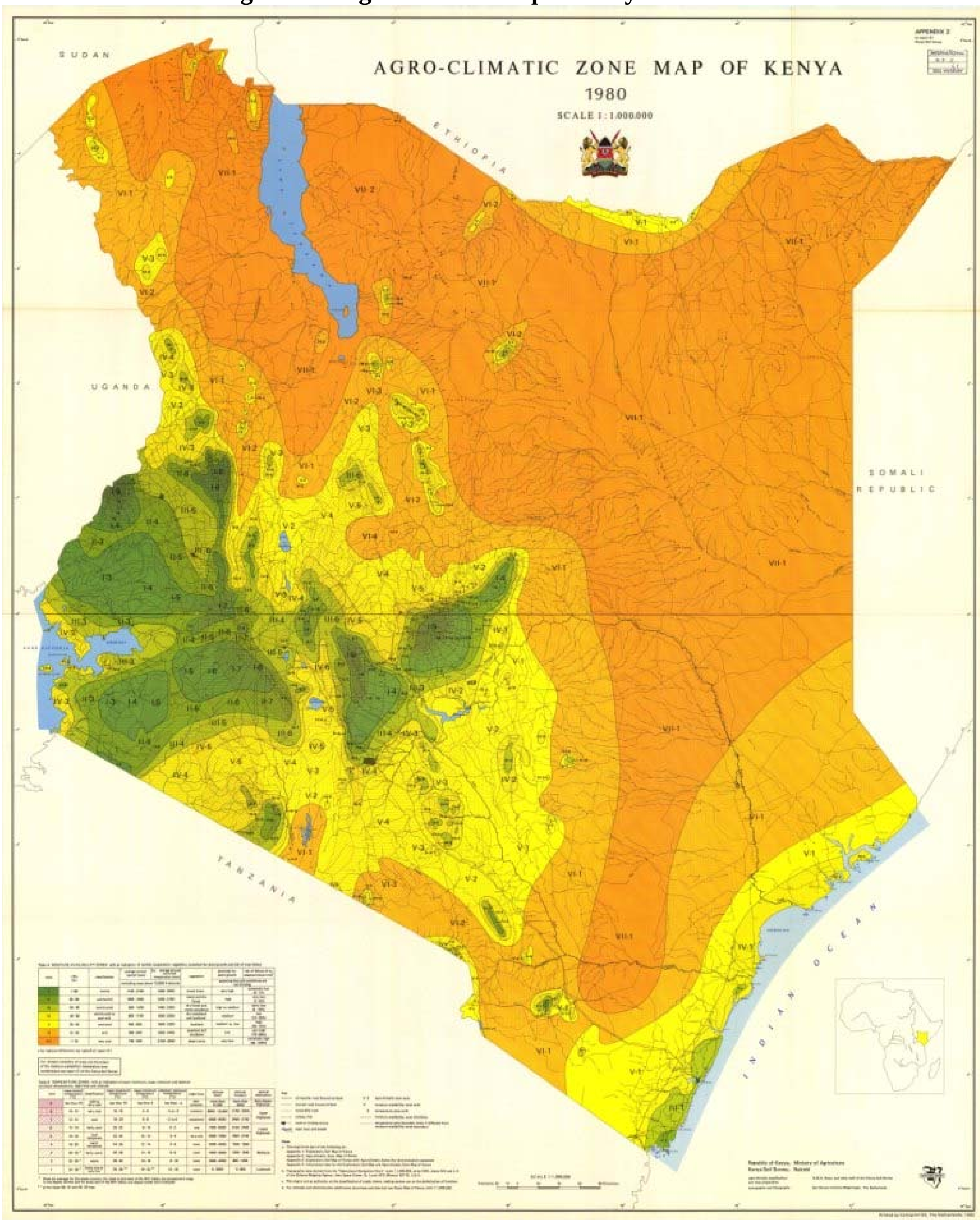
Appendix 12: Maps of Kenya

Figure A.4 Soil map from 1980 survey off of which soil properties are based



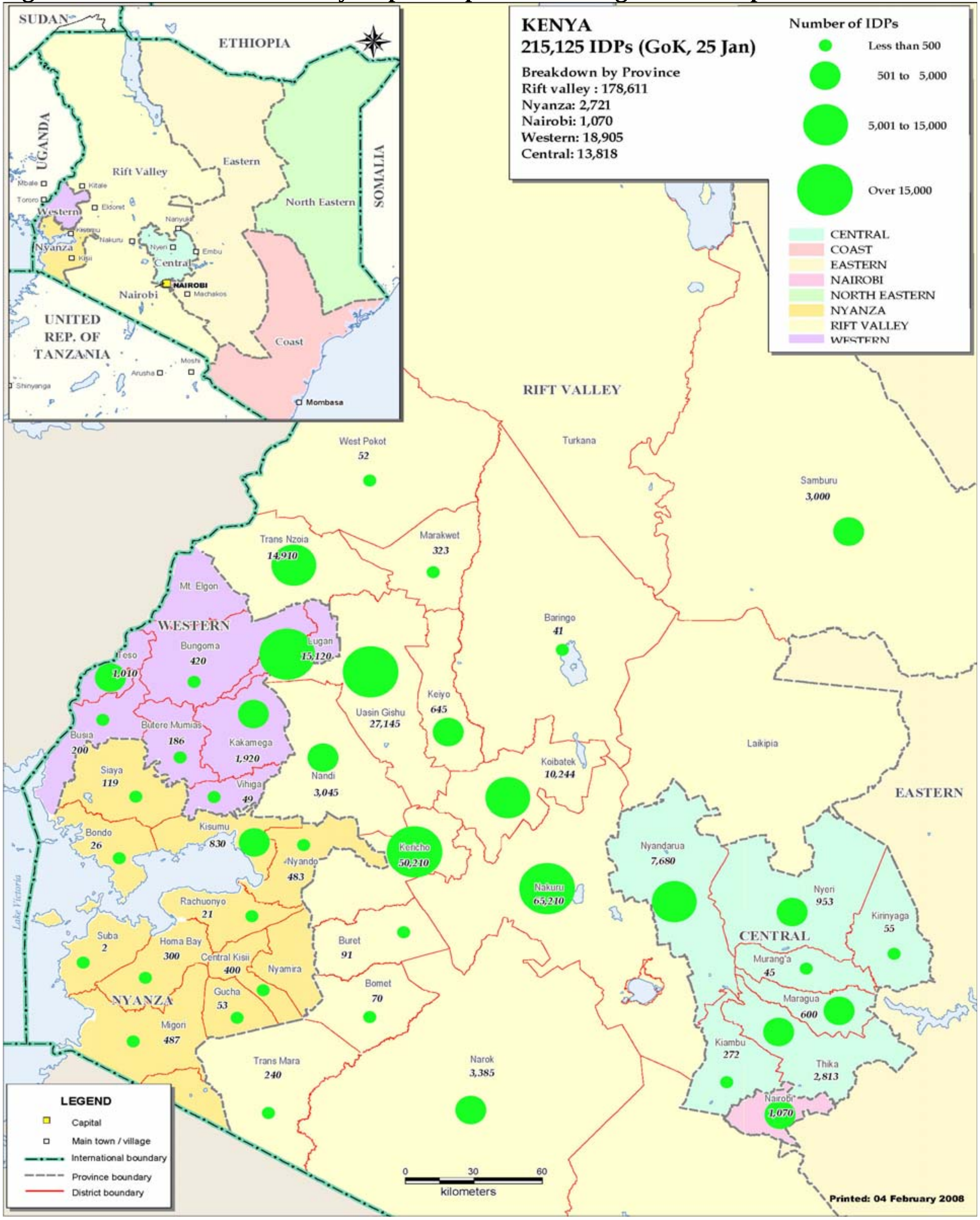
Source: ISRIC – World Soil Information Database (http://library.wur.nl/WebQuery/isric/6336)

Figure A.5 Agro-climatic map of Kenya from 1980



Source: ISRIC – World Soil Information Database (http://library.wur.nl/WebQuery/isric/6336)

Figure A.6 Number of internally displaced persons during 2007-2008 post-election violence



Source: Internal Displacement Monitoring Center (http://www.internal-displacement.org)

REFERENCES

REFERENCES

Abdoulaye, T., & Sanders, J. H. (2005). Stages and determinants of fertilizer use in semiarid African agriculture: the Niger experience. *Agricultural Economics,* 32, 167-179.

Ackello-Ogutu, C., Paris, Q., & Williams, W. A. (1985). Testing Function a Crop Liebig Response against Polynomial Specifications. *American Journal of Agricultural Economics,* 67(4), 873-880.

Ae, N., Arihara, J., Okada, K., Yoshihara, T., & Johansen, C. (1990). Phosphorus Uptake by Pigeon Pea and Its Role in Cropping Systems of the Indian Subcontinent. *Science,* 248(4954), 477-480.

Agbenin, J. O., & Goladi, J. T. (1997). Carbon, nitrogen and phosphorus dynamics under continuous cultivation as influenced by farmyard manure and inorganic fertilizers in the savanna of northern Nigeria. *Agriculture, Ecosystems & Environment*, 63(1), 17-24.

Alene, A. D., Manyong, V., Omanya, G., Mignouna, H., Bokanga, M., & Odhiambo, G. (2008). Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *Food Policy,* 33(4), 318-328.

Anderson, J. R., Dillion, J. L., & Hardaker, B. (1977). *Agricultural Decision Analysis.* Ames, Iowa: The Iowa State University.

Ariga, J. (forthcoming). Essays on Farm Fertilizer Profitability and Demand. Michigan State University. PhD dissertation. Department of Agricultural Economics. Michigan State University. East Lansing.

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 00921. Washington DC.

Ariga, J., Jayne, T. S., & Nyoro, J. K. (2006). Factors Driving the Growth in Fertilizer Consumption in Kenya, 1990-2005: Sustaining the Momentum in Kenya and Lessons for Broader Replicability in Sub-Saharan Africa. Working Paper Series No. 24. Nairobi, Kenya: Tegemeo Institute.

Ariga, J., Jayne, T. S., Kibaara, B., & Nyoro, J. K. (2008). Trends and Patterns in Fertilizer Use By Smallholder Farmers in Kenya, 1997-2007. Working Paper Series No. 28. Nairobi, Kenya: Tegemeo Institute.

Banful, A. B. (2011). Old Problems in the New Solutions? Politically Motivated Allocation of Program Benefits and the “New” Fertilizer Subsidies. *World Development,* 39(7), 1166-1176.

Barrett, C. B. (1996). On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics,* 51, 193-215.

Bationo, A., Christianson, C. B., Baethgen, W. E., & Mokwunye, A. U. (1992). A farm-level evaluation of nitrogen and phosphorus fertilizer use and planting density for pearl millet production in Niger. *Fertilizer Research,* 31(2), 175-184.

Bauer, A., & Black, A. L. (1994). Quantification of the Effect of Soil Organic Matter Content on Soil Productivity. *Soil Science Society of America Journal*, 58(January-February), 185-193.

Bellemare, M. F., & Barrett, C. B. (2006). An Ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia. *American Journal of Agricultural Economics,* 88(2), 324-337.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.*

Berck, P., & Helfand, G. (1990). Reconciling the von Liebig and Differentiable Crop Production Functions. *American Journal of Agricultural Economics,* 72(4), 985-996.

Berck, P., Stohs, S., & Geoghegan, J. (2000). A Strong Test of the von Liebig Hypothesis. *American Journal of Agricultural Economics*, 82(4), 948-955.

Berry, A. R., & Cline, W. R. (1979). *Agrarian structure and productivity in developing countries.* Baltimore, MD: Johns Hopkins University Press.

Bumb, B. (1995). Global Fertilizer Perspective, 1980–2000: The Challenges in Structural Transformation. Muscle Shoals, AL: IFDC.

Byrnes, B. H. (1990). Environmental effects of N fertilizer use - An overview. *Fertilizer Research,* 26(1-3), 209-215.

Chamberlain, G. (1980). Analysis of Covariance Data with Qualitative. *The Review of Economic Studies*, 47(1), 225-238.

Chaplin III, F. S., Vitousek, P. M., & Van Cleve, K. (1986). The Nature of Nutrient Limitation in Plant Communities. *The American Naturalist,* 127(1), 48-58.

Chayanov, A. V. (1962). In D. Thorner (Ed.), *The theory of peasant economy.* Homewood, IL: Irwin.

Christiaensen, L., & Demery, L. (2007). *Down to Earth: Agriculture and Poverty Reduction in Africa.* Washington DC: The World Bank.

Cleaver, K. M., & Schreiber, G. A. (1994). *Reversing the spiral, the population, agriculture and environment nexus in sub-Saharan Africa.* Washington DC: World Bank.

Coady, D. P. (1995). An Empirical Analysis of Fertilizer Use in Pakistan. *Economica,* 62(246), 213-234.

Coase, R. H. (1960). The Problem of Social Cost. *Journal of Law and Economics,* 3, 1-44.

Crawford, E., Kelly, V., Jayne, T.S., & Howard, J. (2003). Input use and market development in Sub-Saharan Africa: an overview. *Food Policy*, 28(4), 277-292.

Croppenstedt, A., Demeke, M., & Meschi, M. M. (2003). Technology Adoption in the Presence of Constraints: the Case of Fertilizer Demand in Ethiopia. *Review of Development Economics,* 7(1), 58-70.

Cunningham, K. M., & Ogilvie, J. C. (1972). Evaluation of hierarchical grouping techniques: a preliminary study. *The Computer Journal*, 15(3), 209-213.

Daramola, B. (1989). The study of socio-economic factors influencing fertilizer adoption decisions in Nigeria: A survey of Oyo State farmers. *Fertilizer Research*, 20(3), 143-151.

Dorward, A. (2009). Rethinking Agricultural Input Subsidy Programmes in a Changing World. Policy. London, UK: Center for Development, Environment and Policy. School of Oriental and African Studies.

Doss, C. R. (2001). Designing Agricultural Technology for African Women Farmers: Lessons from 25 Years of Experience. *World Development*, 29(12), 2075-2092.

Doss, C. R., & Morris, M. L. (2001). How does gender affect the adoption of agricultural innovations? The case of improved maize technology in Ghana. *Agricultural Economics,* 25(1), 27-39.

Drechsel, P., Gyiele, L., Kunze, D., & Cofie, O. (2001). Population density, soil nutrient depletion, and economic growth in sub-Saharan Africa. *Ecological Economics,* 38(2), 251-258.

Duda, R. O., & Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. Chichester, West Sussex, UK: Wiley.

Duflo, E., Kremer, M., & Robinson, J. (2008). How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *American Economic Review*, 98(2), 482-488.

Duflo, E., Kremer, M., & Robinson, J. (2009). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya.

Ellis, F. (1992). *Agricultural Policies in Developing Countries.* New York: Cambridge University Press.

Esipisu, I. (2011, August 22). Lime use improves maize harvest on acidic Kenyan soils. African Agriculture Blog. Retrieved from http://www.africanagricultureblog.com/2011/08/lime-use-improves-maize-harvest-on.html

FAO. (1975). Planning and organization of fertilizer use development in Africa. Rome, Italy.
Feder, G, & Umali, D. L. (1993). The adoption of agricultural innovations: a review. *Technological Forecasting and Social Change*, 43(3-4), 215-239.

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

Feller, C., & Beare, M. (1997). Physical control of soil organic matter dynamics in the tropics. *Geoderma*, 79(1-4), 69-116.

Freeman, H. A., & Omiti, J. M. (2003). Fertilizer use in semi-arid areas of Kenya: analysis of smallholder farmers' adoption behavior under liberalized markets. *Nutrient Cycling in Agroecosystems,* 66, 23-31.

Gebremedhin, B., & Swinton, S. M. (2003). Investment in soil conservation in northern Ethiopia: the role of land tenure security and public programs. *Agricultural Economics,* 29(1), 69-84.

Giller, K. E., & Cadisch, G. (1995). Future benefits from biological nitrogen fixation: An ecological approach to agriculture. *Plant and Soil,* 174(1-2), 255-277.

Giller, K. E., Cadisch, Georg, Ehaliotis, C., & Adams, E. (1997). Building Soil Nitrogen Capital in Africa. In P.A. Sanchez, R.J. Buresh, & F. Calhoun (Eds.), *Replenishing Soil Fertility in Africa* (pp. 151-192). Madison, Wisconsin: Soil Science Society of America.

Griffin, R. C., Montgomery, J. M., & Rister, M. E. (1987). Selecting Functional Form in Production Function Analysis. *Western Journal of Agricultural Economics,* 12(2), 216-227.

Griffith, B. (n.d.). Essential Role of Phosphorus in Plants (pp. 1-8).

Grimm, S. S., Paris, Q., & Williams, W. A. (1987). A von Liebig Model for Water and Nitrogen Crop Response. *Western Journal of Agricultural Economics*, 12(2), 182-192.

Groffman, P. M., Hendrix, P. F., & Crossley, D. A. (1987). Nitrogen dynamics in conventional and no-tillage agroecosystems with inorganic fertilize or legume nitrogen inputs. *Plant and Soil,* 332, 315-332.

Gulati, A, & Narayanan, S. (2003). *The Subsidy Syndrome in Indian Agriculture.* Oxford: Oxford University Press.

Hassan, R. M., Murithi, F., & Kamau, G. (1998). Determinants of fertilizer use and the gap between farmers' maize yields and potential yields in Kenya. In R. Hassan (Ed.), *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya* (pp. 137-161).

Hassan, R. M., Njoroge, K., Njore, M., Otsyula, R., & Laboso, A. (1998). Adoption patterns and performance of improved maize in Kenya. In R.M. Hassan (Ed.), *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya* (pp. 107-136).

Heisey, P. W., & Mwangi, W. (1997). Fertilizer Use and Maize Production. In D. Byerlee & C. K. Eicher (Eds.), *Africa's Emerging Maize Revolution* (pp. 193-211). Boulder, CO: Lynne Rienner Publishers.

Heisey, P., & Norton, G. (2007). Fertilizers and other farm chemicals. In R. Evenson & P. Pingali (Eds.), *Handbook of Agricultural Economics* (A., Vol. 3, pp. 2741-2777).

Hopper, D. (1993). Indian Agriculture and Fertilizer: An Outsider's Observations. New Delhi: Keynote address to the Fertiliser Association of India (FAI) seminar on "Emerging Scenarios in Fertilizer and Agriculture: Global Dimensions."

Huffman, W. E. (1974). Decision Making: The Role of Education. *American Journal of Agricultural Economics, 56*(1), 85-97.

IUSS Working Group WRB. (2007). World reference base for soil resources 2006. Rome, Italy: Food and Agriculture Organization.

Isham, J. (2002). The Effect of Social Capital on Fertiliser Adoption: Evidence from Rural Tanzania. *Journal of African Economies,* 11(1), 39-60.

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

Jayne, T. S., Yamano, T., Nyoro, J., & Awuor, T. (2001). Do farmers really benefit from high food prices? Balancing rural interests in Kenya's maize pricing and marketing policy. Working Paper Series No. 2B. Nairobi, Kenya: Tegemeo Institute.

Johnson, M., Hazell, P., & Gulati, Ashok. (2003). The Role of Intermediate Factor Markets in Asia's Green Revolution: Lessons for Africa? *American Journal of Agricultural Economics,* 85(5), 1211-1216.

Kaliba, A. R. M., Verkuijl, H., & Mwangi, Wilfred. (2000). Factors Affecting Adoption of Improved Maize Seeds and Use of Inorganic Fertilizer for Maize Production in the Intermediate and Lowland Zones of Tanzania. *Journal of Agricultural and Applied Economics,* 32(1), 35-47.

Kapkiyai, J. J., Karanja, N. K., Qureshi, J. N., Smithson, P. C., & Woomer, P. L. (1999). Soil organic matter and nutrient dynamics in a Kenyan nitisol under long-term fertilizer and organic input management. *Soil Biology and Biochemistry,* 31(13), 1773-1782.

Kebede, Y., Gunjal, K., & Coffin, G. (1990). Adoption of New Technologies in Ethiopian Agriculture: The Case of Tegulet-Bulga District, Shoa Province. *Agricultural Economics,* 4, 27-43.

Kelly, V. (2005). Farmers' demand for fertilizer in Sub-Saharan Africa. East Lansing, Michigan.

Key, N., Sadoulet, E., & de Janvry, A. (2000). Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics,* 82(2), 245-259.

Kherallah, M., Delgado, C. L., Gabre-Madhin, E. Z., Minot, N., & Johnson, M. (2002). *Reforming Agricultural Markets in Africa.* Baltimore, MD: Johns Hopkins University Press.

Kimuyu, P., Jama, M., & Muturi, W. (1991). Determinants of fertilizer use on smallholder coffee and maize in Muranga District, Kenya. *Eastern Africa Economic Review,* 7(1), 1-11.

Kironchi, G., Mbuvi, J. P., & Nguluu, S. (2006). Analysis of Climate Data and the Associated Risks to Maize Production in Semi-Arid Eastern Kenya (pp. 115-122).

Komicha, H. H., & Ohlmer, B. (2007). Influence of credit constraint on technical efficiency of farm households in Southeastern Ethiopia.

Kouka, P.-J., Jolly, C., & Henao, J. (1995). Agricultural response functions for limited resource farmers in sub-Saharan Africa. *Fertilizer Research,* (40), 135-141.

Kumbhakar, S. C., & Baushan, B. (2009). Modelling farm production decisions under an expenditure constraint. *European Review of Agricultural Economics*, 1-25.

Lanzer, E. A., & Paris, Q. (1981). A New Analytical Framework for the Fertilization Problem, *American Journal of Agricultural Economics,* 63(1), 93-103.

Larson, B., & Frisvold, G. (1996). Fertilizers to support agricultural development in sub-Saharan Africa: what is needed and why. *Food Policy,* 21(6), 509-525.

Ledgard, S. E., & Stelle, K. W. (1992). Biological nitrogen fixation in mixed legume/grass pastures. In J. K. Ladha, T. George, & B. B. Bohlool (Eds.), *Biological nitrogen fixation for sustainable agriculture* (pp. 137-153). Dordrecht, The Netherlands: Kluwer Academic.

Li, G., Rozelle, S., & Brandt, L. (1998). Tenure, land rights, and farmer investment incentives in China. *Agricultural Economics,* 19, 63-71.

Liu, Y., & Myers, R. (2009). Model selection in stochastic frontier analysis with an application to maize production in Kenya. *Journal of Productivity Analysis,* 31(1), 33-46.

Maobe, S. N., Kidula, N. L., & Ondicho, A. R. (2000). Effect of Green Manure Residue Management Practices on Maize Yield in Southwest Kenya. Kisii, Kenya: Kenya Agricultural Research Institute.

Marenya, P. P., & Barrett, C. B. (2009a). State-conditional Fertilizer Yield Response on Western Kenyan Farms. *American Journal of Agricultural Economics,* 91(4), 991-1006.

Marenya, P. P., & Barrett, C. B. (2009b). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics,* 40(5), 561-572.

Matsumoto, T., & Yamano, T. (2011). Optimal Fertilizer Use on Maize Production in East Africa. In T. Yamano, K. Otsuka, & F. Place (Eds.), *Emerging Development of Agriculture in East Africa: Markets, Soil and Innovations* (pp. 117-132). Springer Netherlands.

Mbata, J. N. (1997). Factors Influencing Fertilizer Adoption and Rates of Use Among Small-Scale Food Crop Farmers in the Rift Valley Area in Kenya. *Quarterly Journal of International Agriculture*, 36, 285-302.

Meertens, B. (2005). A realistic view on increasing fertiliser use in sub-Saharan Africa. Retrieved from www.meertensconsult.nl

Merckx, Roel, Diels, Jan, Vanlauew, B., Sanginga, Nteranya, Denef, K., & Oorts, K. (2001). Soil Organic Matter and Soil Fertility. In G. Tian, F. Ishida, & D. Keatinge (Eds.), *Sustaining Soil Fertility in West Africa* (pp. 69-89). Madison, Wisconsin: Soil Science Society of America.

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika,* 45(3), 325-342.

Ministry of Agriculture. (2008). Economic review of agriculture: 2008. Nairobi, Kenya.

Minot, Nicholas, & Benson, T. (2009). Fertilizer Subsidies in Africa: Are Vouchers the Answer? IFPRI Issue Brief 60. Washington, DC: International Food Policy Research Institute.

Mooney, H. A., Vitousek, P. M., & Matson, P. A. (1987). Exchange of materials between terrestrial ecosystems and the atmosphere. *Science,* 238, 926-932.

Morris, M., Kelly, V., Kopicki, R. J., & Byerlee, Derek. (2007). Fertilizer Use in African Agriculture: Lessons Learned and Good Practice Guidelines. Washington DC: World Bank.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1), 69-85.

Murwira, K. H., Swift, M. J., & Frost, P. G. H. (1995). Manure as a key resource in sustainable agriculture. In J. M. Powell (Ed.), *Livestock and sustainable nutrient cycling in mixed farming systems of sub-Saharan Africa, Vol. III.* Addis Ababa, Ethiopia: Tech. Pap. Int. Livestock Ctr. for Africa.

Muyanga, M. (forthcoming). Impact of Increasing Population Densities on Land Use and Agricultural Commercialization in Kenya: A panel data analysis. PhD dissertation. Department of Agricultural Economics. Michigan State University. East Lansing.

Nkamleu, G. (2000). Determinants of chemical input use in peri-urban lowland systems: bivariate probit analysis in Cameroon. *Agricultural Systems*, 63(2), 111-121.

Obare, G., Omamo, S. W., & Williams, J. (2003). Smallholder production structure and rural roads in Africa: the case of Nakuru District, Kenya. *Agricultural Economics,* 28(3), 245-254.

Odhiambo, J. J. O., & Magandini, V. N. (2008). An assessment of the use of mineral and organic fertilizers by smallholder farmers in Vhembe district, Limpopo province, South Africa. *Journal of Agricultural Research,* 3(5), 357-362.

Olwande, J., Sikei, G., & Mathenge, M. (2009). Agricultural Technology Adoption: A Panel Analysis of Smallholder Farmers' Fertilizer Use in Kenya. Working Paper No. AfD-0908. Center of Evaluation for Global Action. UC Berkeley.

Ouma, J. O., De Grotte, H., & Owuor, G. (2006). Determinants of Improved Maize Seed and Fertilizer use In Kenya: Policy Implications. Contributed paper prepared for presentation at the International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006.

Ovuka, M., & Lindqvist, S. (2000). Rainfall Variability in Murang'a District, Kenya: Meteorological Data and Farmers' Perceptions. Geografiska Annaler. *Series A, Physical Geography,* 82(1), 107-119.

Piha, M. I., & Munns, D. N. (1987). Nitrogen fixation potential of beans (Phaseolus vulgaris L.) compared with other grain legumes under controlled conditions. *Plant and Soil,* 182, 169-182.

Probert, M. E., Okalebo, J. R., & Jones, R. K. (1995). The use of manure on smallholders' farms in semi-arid eastern Kenya. *Experimental Agriculture,* 31, 371-381.

Rao, M. R., & Mathuva, M. N. (2000). Legumes for improving maize yields and income in semi-arid Kenya. *Agriculture, Ecosystems & Environment,* 78(2), 123-137.

Rauniyar, G., & Goode, F. M. (1992). Technology adoption on small farms. *World Development,* 20(2), 275-282.

Renkow, M., Hallstrom, D. G., & Karanja, D. D. (2004). Rural infrastructure, transactions costs and market participation in Kenya. *Journal of Development Economics,* 73(1), 349-367.

Ricker-Gilbert, J., Jayne, T.S., & Chirwa, E. (2011). Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi. *American Journal of Agricultural Economics*, 93(1), 26-42.

Sanchez, Pedro A., Shepherd, K. D., Soule, M. J., Place, F. M., Buresh, Roland J., & Izac, A.-M. N. (1997). Soil Fertility Replenishment in Africa: An Investment in Natural Resource Capital. In P.A. Sanchez, R.J. Buresh, & F. Calhoun (Eds.), *Replenishing Soil Fertility in Africa* (pp. 1-46). Madison, Wisconsin: Soil Science Society of America.

Sanchez, P.A., Buresh, R.J., & Calhoun, F. (1997). *Replenishing Soil Fertility in Africa*. Madison, Wisconsin: Soil Science Society of America.

Sauer, J., & Tchale, H. (2009). The Economics of Soil Fertility Management in Malawi. *Review of Agricultural Economics,* 31(3), 535-560.
Sen, A. (1962). An Aspect of Indian Agriculture. Economics Weekly.

Shaviv, A., & Mikkelsen, R. L. (1993). Controlled-release fertilizers to increase efficiency of nutrient use and minimize environmental degradation: A review. *Fertilizer Research,* 35(1-2), 1-12.

Sileshi, G., Akinnifesi, F. K., Debusho, L. K., Beedy, T., Ajayi, O. C., & Mong'omba, S. (2010). Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. *Field Crops Research,* 116(1-2), 1-13.

Smaling, E., Nandwa, S., Prestele, H., Roetter, R., & Muchena, F. (1992). Yield response of maize to fertilizers and manure under different agro-ecological conditions in Kenya. *Agriculture, Ecosystems & Environment,* 41(3-4), 241-252.

Smillie, J., & Gershuny, G. (1999). *The Soul of Soil (Fourth Edition*). White River Junction, Vermont: Chelsea Green Publishing Company.

Snapp, S. S. (1998). Soil Nutrient Status of Smallholder Farms in Malawi. *Communications in Soil Science and Plant Analysis,* 27(17 & 18), 2571-2588.

Snapp, S. S., Blackie, M. J., & Donovan, C. (2003). Realigning research and extension to focus on farmers' constraints and opportunities. *Food Policy,* 28(4), 349-363.

Snapp, S. S., Blackie, M. J., Gilbert, R. A, Bezner-Kerr, R., & Kanyama-Phiri, G. Y. (2010). Biodiversity can support a greener revolution in Africa. Proceedings of the National Academy of Sciences.

Stoorvogel, J. J., & Smaling, E. (1990). *Assessment of soil nutrient depletion in sub-Saharan Africa: 1983-2000.* Wageningen, the Netherlands: Winand Staring Ctr.

Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica,* 79(1), 159-209.

Tabu, I. M., Bationo, A., Obura, R. K., & Masinde, J. K. (2007). Effect of Rock Phosphate, Lime and Green Manure on Growth and Yield of Maize in a Non Productive Niche of a Rhodic

Ferralsol in Farmer's Fields. In A. Bationo (Ed.), *Advances in Integrated Soil Fertility Management in Sub-Saharan Africa: Challenges and Opportunities* (pp. 449-456).

The Economist. (2009). Mobile marvels: A special report on telecoms in emerging markets. September 26th 2009 edition, 1-14.

Tittonell, P., Vanlauwe, B., Leffelaar, P. A., Roe, E. C., & Giller, K.E. (2005). Exploring diversity in soil fertility management of smallholder farms in western Kenya: Heterogeneity at region and farm scale. *Agriculture, Ecosystems & Environment,* 110(3-4), 149-165.

Tomich, T. P., Kilby, P., & Johnston, B. F. (1995). *Transforming Agrarian Economies: Opportunities Seized, Opportunities Missed.* New York: Cornell University Press.

Traxler, G., & Byerlee, Derek. (1993). Joint-Product Analysis of the Adoption of Modern Cereal Varieties in Developing Countries. *American Journal of Agricultural Economics,* 75(4), 981-989.

UNHCR. (2008). Report from OHCHR Fact-finding Mission to Kenya, 6-28 February 2008. Geneva, Switzerland.

Vanlauwe, B., Aihou, K., Aman, S., Tossah, B. K., Diels, J., Lyasse, O., Hauser, S., et al. (2000). Nitrogen and phosphorus uptake by maize as affected by particulate organic matter quality, soil characteristics, and land-use history for soils from the West African moist savanna zone. *Biology and Fertility of Soils,* 30(5-6), 440-449.

Vansambeek, J., Ponderjr, F., & Rietveld, W. (1986). Legumes increase growth and alter foliar nutrient levels of black walnut saplings. *Forest Ecology and Management,* 17(2-3), 159-167.

Waithaka, M. M., Thornton, P. K., Shepherd, K. D., & Ndiwa, N. N. (2007). Factors affecting the use of fertilizers and manure by smallholders: the case of Vihiga, western Kenya. *Nutrient Cycling in Agroecosystems,* 78(3), 211-224.

Weight, D., & Kelly, V. (1999). Fertilizer Impacts on Soil and Crops of Sub-Saharan Africa. Agricultural Economics. MSU International Development Paper No. 21. East Lansing, Michigan.

Williamson, O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law and Economics*, 22(2), 233-261.

Winter-Nelson, A., & Temu, A. (2005). Impacts of prices and transactions costs on input usage in a liberalizing economy: evidence from Tanzanian coffee growers. *Agricultural Economics,* 33(3), 243-253.

Wooldridge, J. M. (2009a). *Introductory Econometrics: A Modern Approach (4th edition*). Mason, Ohio: South-Western Cengage Learning.

Wooldridge, J. M. (2009b). Correlated Random Effects Models with Unbalanced Panels. East Lansing, Michigan.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data (Second Edition)*. Cambridge, Massachusetts: The MIT Press.

Xu, Z., Burke, W. J., Jayne, T.S., & Govereh, J. (2009). Do input subsidy programs "crowd in" or "crowd out" commercial market development? Modeling fertilizer demand in a two-channel marketing system. *Agricultural Economics,* 40(1), 79-94.

Xu, Z., Guan, Z., Jayne, T.S., & Black, R. (2009). Factors influencing the profitability of fertilizer use on maize in Zambia. *Agricultural Economics,* 40(4), 437-446.

Yanggen, D., Kelly, V., Reardon, T., & Naseem, A. (1998). Incentives for Fertilizer Use in Sub-Saharan Africa: A Review of Empirical Evidence on Fertilizer. MSU International Development Working Paper No. 70. East Lansing, Michigan.

Yesuf, M., Mekonnen, A., Kassie, M., & Pender, J. (2005). Cost of Land Degradation in Ethiopia: A Critical Review of Past Studies.