

ESSAYS ON FARM FERTILIZER PROFITABILITY AND DEMAND

By

Joshua Makori Ariga

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Agricultural Economics

2013

ABSTRACT

ESSAYS ON FARM FERTILIZER PROFITABILITY AND DEMAND

By

Joshua Makori Ariga

The 2007/08 increase in world input and output prices put pressure on governments to intervene in markets using various policies including subsidies in an effort to raise agricultural production, incomes, and alleviate poverty and food insecurity. Countries like Russia and China implemented protectionist policies involving export restrictions on fertilizers and cereal outputs in a bid to encourage domestic production and safeguard against high food prices from speculation in futures markets. Such fears also influenced developing countries to subsidize inputs and implement social safety-net programs.

Due to the increased interest in agricultural intensification, it is important for policy makers to be informed on the contribution of fertilizer to farm incomes in different agro-ecological zones so that interventions are tailored to local conditions. Essay 1 uses rigorous econometric methods on a rural household panel dataset to provide insights on the spatial heterogeneity of the effect of fertilizer on yields and household incomes and so the need for location-specific intervention. The results show that using a complementary set of improved technologies (fertilizer and hybrid seed) has significant yield effect. However, under moisture stress conditions, yields are negatively affected for hybrid compared to non-hybrid seed, indicating the importance of using improved technology that is appropriate to specific local conditions.

The results show that it is not profitable to use fertilizers in some zones. There is spatial heterogeneity in Marginal Value-to-Cost Ratio (MVCR) and Average Value-to-Cost Ratio (AVCR) estimates. This has implications on government intervention through blanket non-targeted subsidies that do not take into account the local conditions and profitability of using fertilizers. This is an important contribution that can aid subsidy and other agricultural investment efforts in Kenya. For areas facing uncertain weather conditions, policies that aim to encourage fertilizer use have to tackle the production risks.

Essay 2 explains results from Essay 1 that show differences in demand even within areas where fertilizer is potentially profitable to use. Essay 2 uses econometric approaches that mitigate bias from endogeneity to analyze factors that influence farmers' decision to use fertilizer. Distance to fertilizer seller is shorter, prices lower, and fertilizer use higher in areas with relatively more rainfall and less moisture stress. There is a complementarity between investments in access to information (extension), other infrastructure, and fertilizer adoption. Indicators of wealth like land size, value of agricultural assets, and using tractor or animal draught for land preparation have a positive and significant effect on the probability of purchasing fertilizer, while higher fertilizer prices have negative effect on use.

Therefore, government policy that encourages private investments in the distribution of fertilizers coupled with training on the agronomic aspects and benefits of using fertilizers can be important in raising production. In cases where resources are constrained and fertilizer prices relatively high (as in remote dry areas in low potential regions), intervention in form of targeted subsidies may contribute to adoption of fertilizers.

**Copyright by
JOSHUA MAKORI ARIGA
2013**

To my parents,
Mzee Jackson Ariga Ombonyo and Mary Sikweya
and my wife and children,
Dinah, Luther, and Yvonne Makori

ACKNOWLEDGEMENTS

Let me extend my gratitude to the dissertation committee for its help in finalizing this project. The Chair of the committee, Thomas Jayne, has worn many hats in our interactions. He has been a mentor, colleague, teacher, and above all a friend. His work ethic is impressive and so is his willingness to go out of his way to listen and assist. I cannot thank him enough. My deep appreciation goes to Roy Black who spent countless hours with me working on theory, empirics, and STATA commands. His encouragement, experience dealing with large datasets, and willingness to look at my work-in-progress and get back to me quickly made this work possible. I am grateful to Robert Myers for invaluable techniques I learned from AEC845 and also for his incisive comments and suggestions that were direct and to the point on issues ranging from modeling to the interpretation of the results. Suggestions by Eric Crawford brought order to ideas that were spread haphazardly and made the document more coherent. Jeffrey Andresen provided insights on the role of climatic factors and he was also instrumental in my training on DSSAT software which combines crop, soil and weather data to simulate and compare outcomes from crop management strategies with observed results. Clearly, my committee epitomized synergism.

I got support from other faculty and staff of the Department of Agricultural Economics in the course of my graduate studies. I am especially thankful to Scott Loveridge for enabling a Dissertation Completion Fund when normal funding fizzled out for me, Margaret Beaver for help with data management issues, and Debbie Conway. I am grateful to my friends and fellow students Andrew Muganga, Athur Mabiso, Bill Burke, Mary Mathenge, Sarma Aralas, Jacob Gilbert, Vandana Yadav, Nicole Mason, Epi Katjiuongua, Feng Song, Helder Zavale, Milu Muyanga, Marcus Coleman, Miltone Ayieko, Pandey Vivek, and Ramzi Adjao.

I would like to thank my brothers and sisters: Stephen, Thomas, Josephine, Rachel, Charles, Samuel, Milka, Grace, Jackline, and David for their support and understanding. I give special thanks to my parents, Mzee Jackson Ariga Ombonyo and Mary Sikweya, for their unconditional love and encouragement.

This effort could not have matured were it not for my wife and children who stood by me through some tough times as I juggled my responsibilities as a husband, father, and student. I hope I will repay Dinah, Luther, and Yvonne for their love and support. Thank you.

PREFACE

P.1 General Background: Market Reforms and their Implications for Fertilizer Profitability and Demand

This background covers material that is useful in putting the two essays in the context of policy reforms that paved the way for increased private sector participation and supportive public investments. We provide a summary of Kenya's maize and fertilizer market reforms from a controlled to a free market economy starting in the mid-1990s followed by a mix of public and private investments in infrastructure and services. Up until the early 1990s, the government of Kenya determined the price of maize at the farm, the buying and selling prices applying to millers and retailers, as well as the retail price of maize meal to consumers. These controlled prices were pan-territorial and pan-seasonal, adjusted once every year at the beginning of the planting season. The government marketing board, National Cereals and Produce Board (NCPB), had a longstanding monopoly on internal and external trade and informal private trade across district boundaries was illegal, as was cross-border trade. Traders were required to apply for movement permits to allow them to transport grain across district boundaries and risked jail and fines if they were caught trading in maize or inputs like fertilizers without authorization from relevant state organs. Fertilizer and maize output markets were basically run by state agencies.

By early 1990s the pressure on government budgets to run these agencies was exacerbated by corruption by those charged with overseeing the importation and distribution of inputs and the purchase and sale of maize. This, coupled with the demand by international development partners on transparency, pushed the government to start liberalizing these markets in a piecemeal manner. The reform process intensified in late 1993, when, under pressure from international lenders, the government eliminated movement and price controls on maize trading,

deregulated maize and maize meal prices, and eliminated direct subsidies on maize sold to registered millers (Jayne and Argwings-Kodhek, 1997). By 1995, private traders were officially allowed to transport maize across districts without hindrance. Starting in the 1995/96 marketing year, and under pressure from external donors, the government dramatically reduced the NCPB's operating budget paving the way for increased private trade investments in importation, distribution, and retailing.

The ensuing period witnessed the doubling of national fertilizer consumption from 200000 metric tons in 1990/91 to over 400000 in 2007/08 (Ministry of Agriculture Annual Report, 2008) as shown in Figure 1. The decline in fertilizer imports following the 2006/07 season, as depicted in Figure 1, corresponds to the spike in world prices shown in Figure 2 and civil unrest from a disputed presidential election which disrupted farm activities in some areas. The rise in world prices resulted from increased competition for fertilizer inputs from bio-fuel producers, slow expansion in world fertilizer manufacturing capacity, and increases in petroleum products which are a major ingredient in fertilizer manufacture. This rise in prices partly explains the drop in national imports during this period (Figure 1). However, despite the rise in prices, marketing margins have been declining over time, suggestive of a competitive private sector and / or reduction in transport and transaction costs (Figure 2).

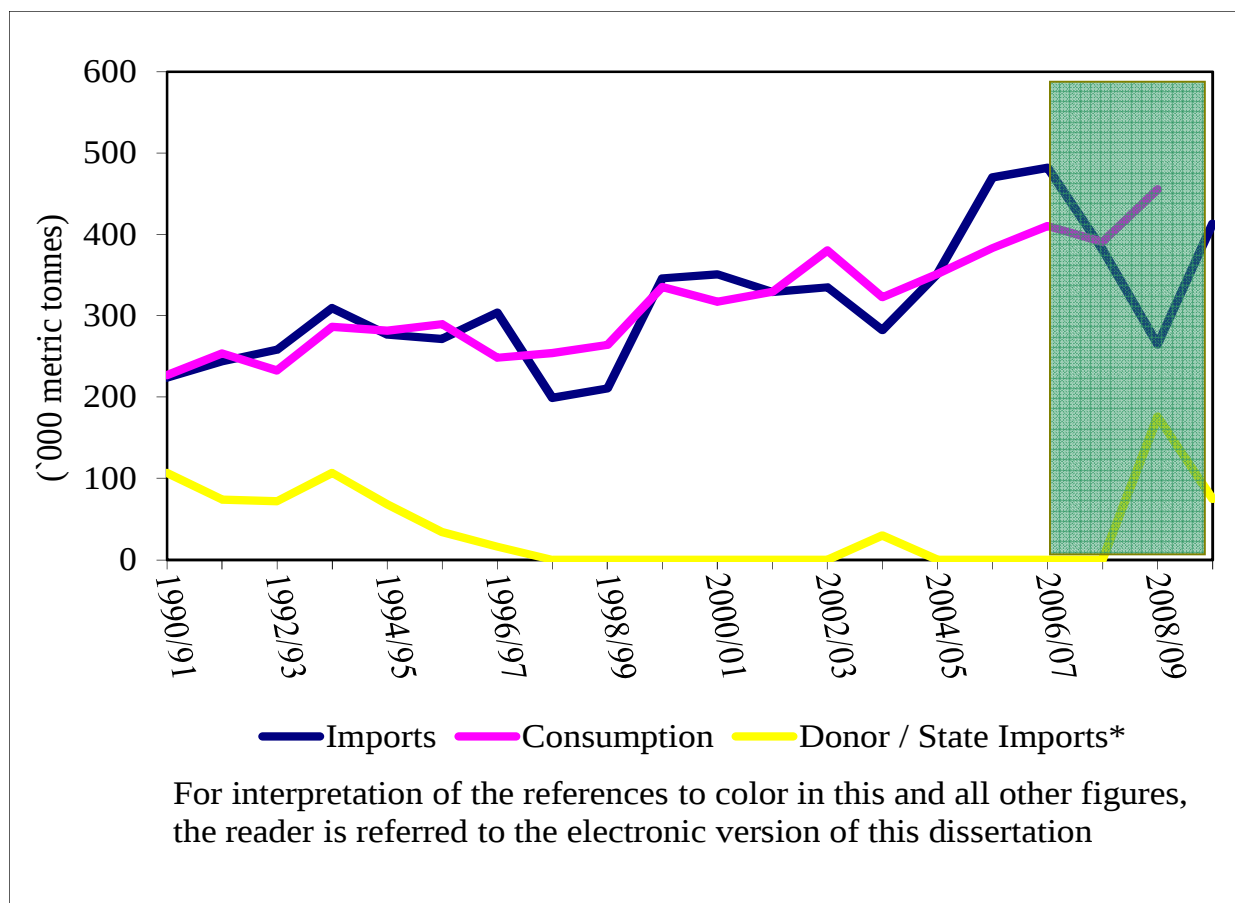


Figure 1. Trends in fertilizer consumption, commercial imports, and donor imports, 1990-2007

Source: Estimated from Ministry of Agriculture (MoA) data in Ariga and Jayne(2010): In 2004 and 2008 NCPB imported approximately thirty and forty percent of national needs (MoA). The estimates for year 2010 are projections for both private and government imports. The years under the color-box cover the time period after 2006/07 when government imports / subsidies re-started partly as a reaction to deficits in maize production and post-election violence disruptions of agricultural activities (this period is not covered in detail in this study).

The margin between wholesale world prices (cif, ex-Mombasa port on the east coast) and inland town of Nakuru has been declining over the period covered by this study (1997-2007). The world price was fairly constant over this period but rose sharply after 2006/07. This implies that marketing costs declined leading to lower prices at Nakuru. Studies (Kimuyu 1994; Wanzala et al 2002; Allgood and Kilungo 1996; IFDC 2001) and interviews with stakeholders

suggest this reduction is a result of increased competition after the 1990s reforms, economies of scope resulting from mergers, and access to competitive credit from international sources.

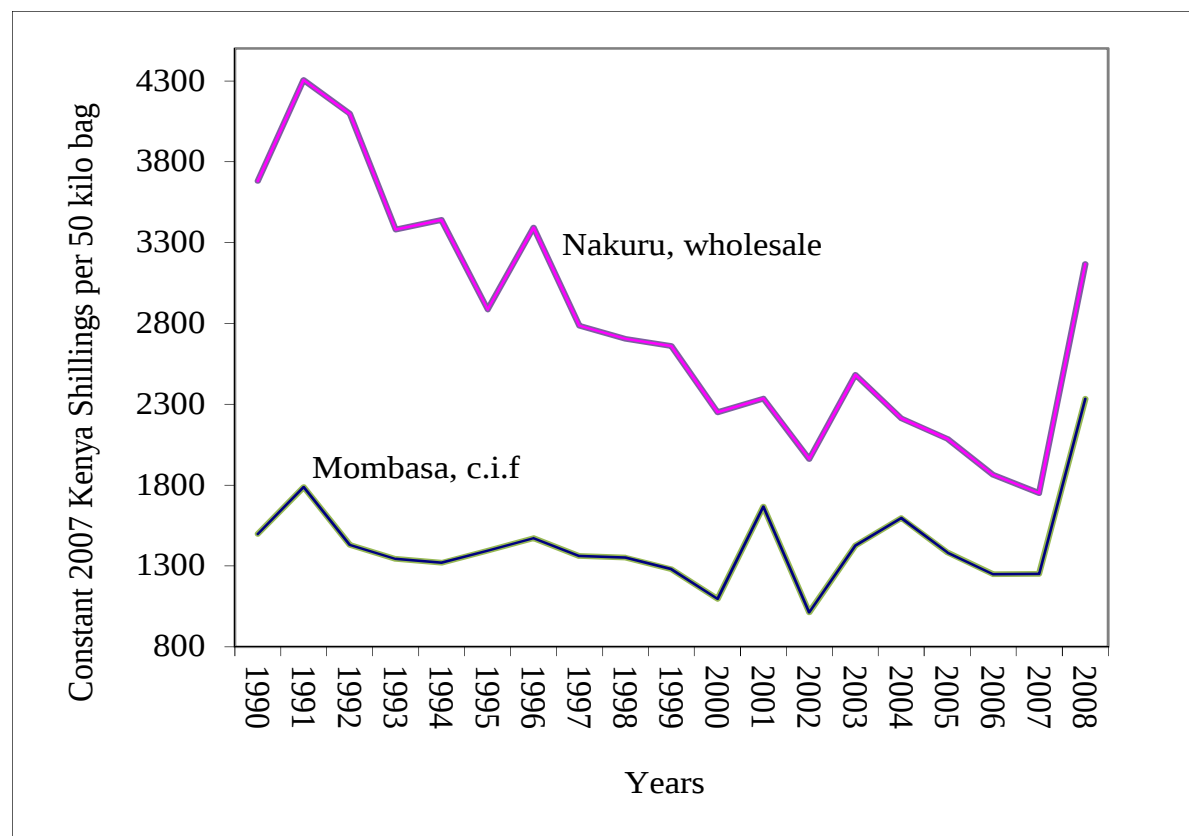


Figure 2. Price of Diammonium Phosphate (DAP) in Mombasa and Nakuru (constant 2007 Kenyan shillings per 50-kg bag)

Source: Ministry of Agriculture. FMB weekly fertilizer reports for c.i.f. Mombasa. This figure was extracted from Ariga and Jayne (2010)

Accompanying the liberalization of input and output markets was the expansion of public and private sector investments in goods and services across the country. Tegemeo Institute¹ household panel data reveals significant declines in distances from farm to tarmac roads, veterinary services, clean water, and electricity (public investments), and distances from farm to

¹ These surveys were funded by USAID and managed jointly by Egerton University (Kenya) and Michigan State University (USA)

the nearest fertilizer seller (private and public investments) as shown in Figure 11 and Table 11 in Appendix 2. There has been an increased investment in private trade in fertilizer and maize markets in response to incentives resulting from de-controlling the prices and the ensuing arbitrage opportunities. So places with relatively high potential for increasing agricultural productivity but which had thin markets under a government-run regime started getting increased private trade activity after liberalization. Comparing the period from the implementation of these reforms (mid 1990s) to 2007 there was a noticeable increase in fertilizer use and yields on smallholder maize plots (Tegemeo Institute Household Survey Reports).

Figure 3 provides a schematic depiction of how public investments in market infrastructure and policy reform of the fertilizer and maize markets generated a number of responses from the private sector, leading to changes in smallholder farm behavior. There are some synergies between liberalization of input and maize markets and public investments in support of smallholder agriculture, leading to substantial private sector investment in fertilizer retailing and maize marketing, which in turn resulted in an increase in fertilizer use and yields on smallholder maize farms.

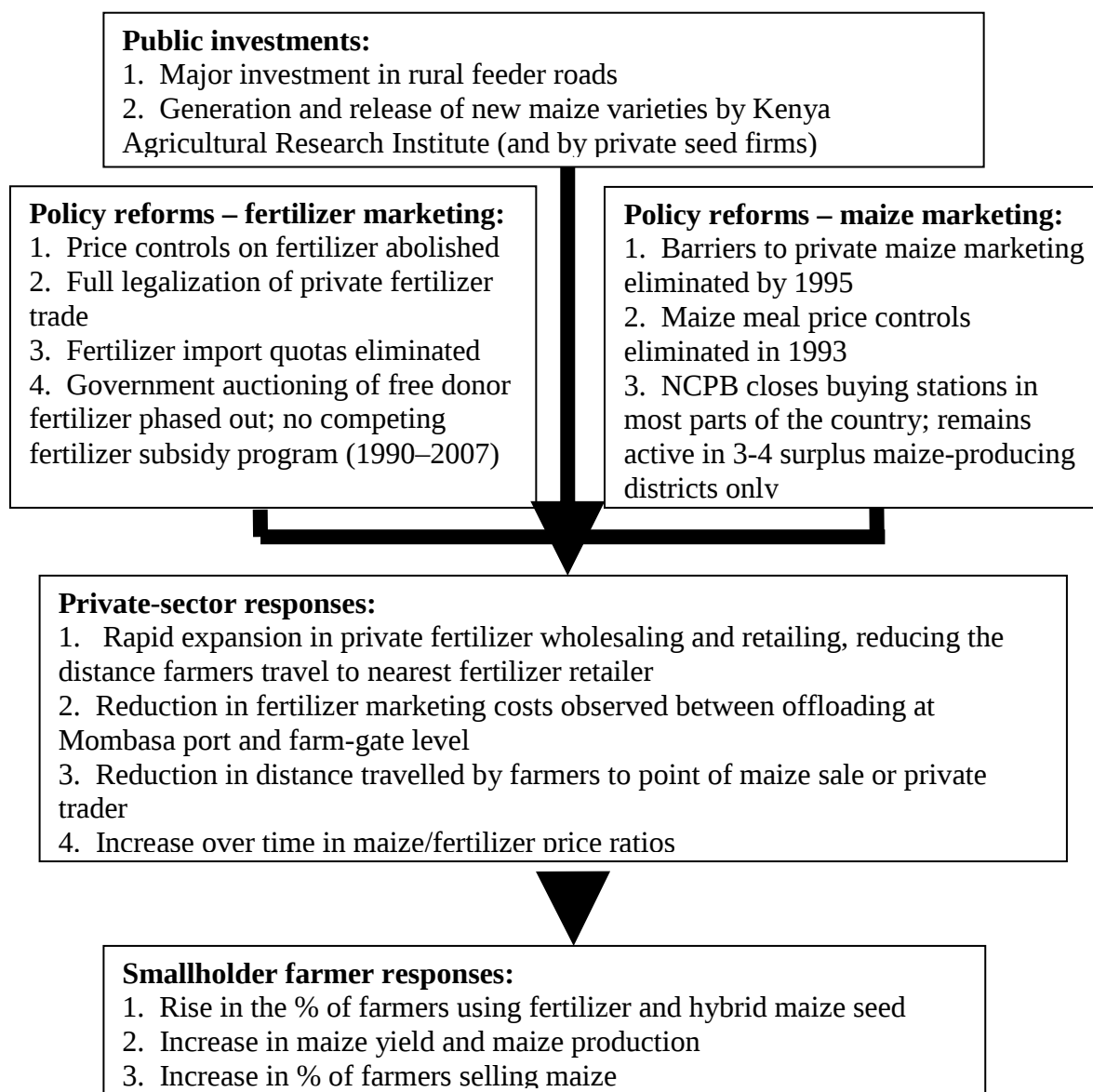


Figure 3. Synergies between public goods investments, policies, and private-sector response in promoting fertilizer use and maize yield improvements by smallholder farmers

P.2 General Description of the Data

This section provides a broad look at the data (detailed description is given Appendix A and B. The data used in the analysis comes from Tegemeo Rural Household Survey for the years

1997, 2000, 2004, 2007. This survey was funded by USAID and implemented jointly by Tegemeo Institute (Egerton University, Kenya) and Michigan State University. We focus on plots that are planted with maize because it is a major food crop grown by over 95% the sample households and accounting for over 50% of fertilizer use (Ministry of Agriculture Annual Reports, Tegemeo Household Survey). The data covers mostly production data for major crops (inputs, outputs, and price data) and off-farm activities.

There are two planting seasons, major and minor, with the major season accounting for over 80% of annual maize output. Most households have two maize crops in a year, one during the longer rain period and the other during the shorter season. Short season rainfall maize seed is composed of quick-maturing varieties compared to long season seed types.

In addition to the seasonal nature of production there are two distinct geographic regions with different agro-ecological conditions². The low potential region covers the lowland zones in the eastern and coastal areas while the high potential region consists of the mid and highlands areas in central, Rift valley, and western parts of the country. The latter region has more rainfall, better soils for maize production, and more investment in roads, electricity, and schools probably due to its agricultural potential. This is the region that had the most European settlers during the colonial period and so benefited from inordinate state largesse in development funds. Some of these factors might also explain the concentration of fertilizer retailers in this region.

Therefore, we use long season maize plots for the high potential region for both Essay 1 and 2. Though emphasis is laid on high potential region, results are compared with those from

² Low and high potential regions are designated loosely based on the agro-ecological conditions and the crop yield potential. Districts with relatively higher rainfall, good soils, and better infrastructure are classified under high potential region.

the low potential regions whenever it is relevant. Figure 4 shows the spatial location of villages covered on a map of Kenya.

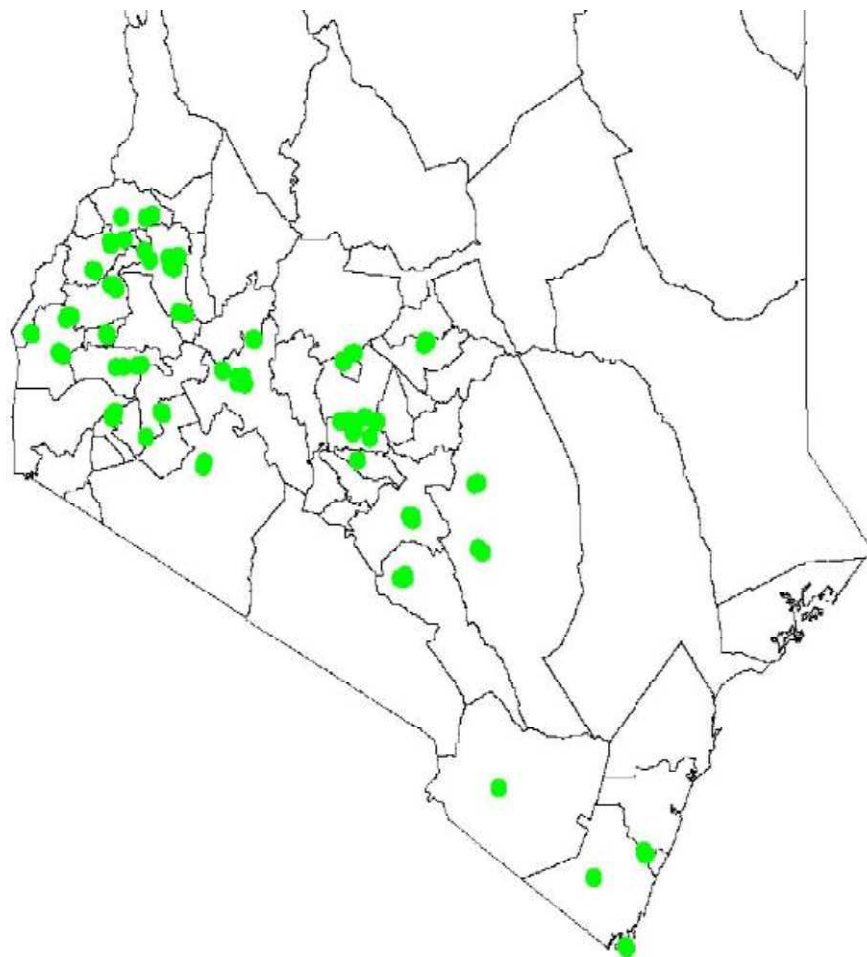


Figure 4. Location of Survey Villages on Map of Kenya

Note: Source: extracted from Suri(2005)

TABLE OF CONTENTS

LIST OF TABLES	xvii
LIST OF FIGURES	xix
ESSAY 1:	
PROFITABILITY OF FERTILIZER: PANEL ESTIMATION OF SMALLHOLDER YIELD RESPONSE IN KENYA	1
1.1.1 Background and Literature Review	1
1.1.2 Objectives	4
1.1.3 Functional Form and Econometric Considerations	5
1.1.3.1 Data and Variable Description	9
1.1.3.2 Generation of Yield Index for Mixed Plots	10
1.1.3.3 Empirical Model	11
1.2.0 Results and Recommendations	13
1.2.1 Effect of Nitrogen Application Rates on Yields and Household income	15
1.2.2 Effect of Gender on Yields	21
1.2.3 Effect of Agro-Ecological Zones on Yields	22
1.2.4 Effect of Seed on yield Index	23
1.3.0 Conclusions and Recommendations	26
APPENDIX	28
APPENDIX 1	29
A1.1: Regression Results and Analysis	29
A1.2: Additional Data Analysis	34
A1.2.1 Agro-zones, Rainfall, and Soil Data	36
BIBLIOGRAPHY	40
BIBLIOGRAPHY 1	41
ESSAY 2:	
ESTIMATING DEMAND FOR FERTILIZER: INCORPORATING ENDOGENEITY OF PRIVATE INVESTMENT IN RETAIL MARKETS	44
2.1.2 Introduction: Importance of Estimating Fertilizer Demand and Profitability	44
2.1.3 Objectives	47
2.1.4 Modeling Framework	49
2.1.4.1 Double Hurdle and Correlated Random Effects	50
2.1.4.2 Modeling Unobserved Heterogeneity	53
2.1.4.3 Testing for Endogeneity of Distance to Fertilizer Retail Services	55
2.2 Empirical Model	57
2.3 Data Description and Methodology	59
2.3.1 Definition and Summary Statistics of Variables	60
2.4 Results	63

2.4.1	Factors Influencing Distance to Fertilizer Seller	63
2.4.2	Estimating Demand for Fertilizer in the Context of Endogeneity	65
2.5	Conclusions and Recommendations	70
APPENDIX		72
APPENDIX 2		73
BIBLIOGRAPHY		94
BIBLIOGRAPHY 2		95

LIST OF TABLES

Table 1. Result of Variance Decomposition (between, within and total variations)	8
Table 2. Description of Variables used in Production Function Regression	9
Table 3. Marginal Products of Inputs Averaged across Farms from Correlated RE and RE Regressions for Continuous Variables (yield index as dependent variable).....	14
Table 4. Results from CRE and RE Regressions for High and Low Potential Regions.....	29
Table 5. Marginal and Average Cost Ratios and Incomes.....	33
Table 6. Description of Variables used in Production Function Regression	34
Table 7. Means of Variables used in Production Function (1997-2007).....	35
Table 8. High Potential Region: Correlation Matrix for Continuous Variables.....	36
Table 9. Low Potential Region: Correlation Matrix for Continuous Variables.....	36
Table 10. Depth and Drainage Characteristics of Soil Types	38
Table 11. Changes in Distance from Farm to the Nearest Fertilizer Seller	73
Table 12. Definition of Variables: Fertilizer Demand.....	74
Table 13. High Potential Region: Descriptive Statistics for Continuous Variables	75
Table 14. Low Potential Region: Descriptive Statistics for Continuous Variables.....	75
Table 15. Descriptive Statistics for Categorical Variables	76
Table 16. High Potential Region: Correlation Matrix for Variables	78
Table 17. Low Potential Region: Correlation Matrix for Variables	79
Table 18. Gross Prices for Hybrid Seed, Grain, and DAP Fertilizer (Shs / kg)	80
Table 19. Indexed Prices for Hybrid Seed, Grain, and DAP Fertilizer (Shs / kg).....	80
Table 20 . Factors Influencing Distance to Nearest Fertilizer Seller	81

Table 21 . Average Marginal Effects for Factors Influencing the Probability of Purchasing Fertilizer.....	83
Table 22. Probit Regression for Factors Influencing the Probability of Purchasing Fertilizer.....	85
Table 23. Marginal Effects for Factors Influencing Quantity Applied.....	88
Table 24. Regression Estimates for Factors Influencing Quantity Applied	91

LIST OF FIGURES

Figure 1. Trends in fertilizer consumption, commercial imports, and donor imports, 1990-2007	ix
Figure 2. Price of Diammonium Phosphate (DAP) in Mombasa and Nakuru (constant 2007 Kenyan shillings per 50-kg bag)	x
Figure 3. Synergies between public goods investments, policies, and private-sector response in promoting fertilizer use and maize yield improvements by smallholder farmers	xii
Figure 4. Location of Survey Villages on Map of Kenya.....	xiv
Figure 5. Predicted Yield versus Nitrogen.....	16
Figure 6. Marginal Value Cost Ratio for Nitrogen by Agro-Zone	17
Figure 7. Household Net Profit Added from Nitrogen Use (USD)	19
Figure 8. Maize Yield by Gender of Head of Household for the High Potential Region.....	21
Figure 9. Maize Yield by Agro-Ecological Zone within the High Potential Area	22
Figure 10. Quantity of maize seed as number of crops planted in plot increases.....	34
Figure 11. Relative changes in indicators of access to markets and services, indexed to 1997 ...	73

ESSAY 1:

PROFITABILITY OF FERTILIZER: PANEL ESTIMATION OF SMALLHOLDER YIELD
RESPONSE IN KENYA

1.1.1 Background and Literature Review

Following the global price rise in 2007/08 and faced with an economy on the slide and declining agricultural productivity, the Kenya government returned to participate in markets in an attempt to alleviate food insecurity. For the mostly agrarian economy, low income levels and increasing food insecurity galvanized international and local efforts aimed at making food available and accessible or offering subsidized farm inputs (particularly fertilizer and seeds) to the poor in order to spur increased production. For instance, the government of Kenya provided free fertilizer and seeds to approximately half a million poor households in 2008 (Kenya Broadcasting Corporation, 2010) through the National Accelerated Agricultural Input Access Program (NAAIAP). Declining effective demand for food and inputs (Crawford et al 2003), low production, poor rainfall and increasingly uncertain weather patterns coupled with underdeveloped infrastructure has exacerbated poverty levels in developing economies like Kenya's (World Bank Country Reports).

This heightened activity was directed at raising yields for food crops as one way of reducing food insecurity. The Kenya government's recent food policy has concentrated in raising maize yield due to its role as a staple food in Kenya. It is noteworthy that this policy re-visited earlier state-led interventions in output and input markets that were deemed inefficient and abandoned in the early 1990s (explained more fully under Section P.1 in the preface). The early 1990s saw inefficient state-run agencies in both maize output and input markets being replaced by private sector traders in a move that liberalized the maize market. Though previous

government policies articulated in various documents - Poverty Reduction Strategy Papers (PRSP), Economic Recovery Strategy (ERS), and Session Papers – were directed at poverty reduction, increasing output, and reducing hunger, Kenya's average maize yields (range 0.5-3.0 ton per hectare) trend well below the potential of 6 tonnes per hectare. It is therefore not surprising that high global prices for food stuffs lead developing countries to increase public spending to raise food supply either through input subsidies or food programs for the poor.

Given this situation where actions by the state are often haphazard and ad-hoc with unpredictable shelf-life it is important to provide policy makers and other key players with information on factors that influence maize yield. In this way, incentives will be directed at areas where they will have most impact and public resources better utilized. In addition, the inter-play between private and public investments in the agricultural sector need to be carefully calibrated in order to avoid the latter “crowding out” the former (Jayne et al 2003) and re-creating the unsustainable environment of the pre-1990 era when government bureaucracies run markets to the exclusion of the private sector. Though there is no clear cut demarcation for the appropriate extent of public participation in markets, it is important to provide policy makers with information on areas that might benefit from the state playing a more active role. Instead of acting without more complete information, public actions can be aided with empirical evidence, leading to targeted interventions.

In this Essay we estimate the effect of various inputs on maize yield, measure the contribution of fertilizer to incomes of smallholder producers in Kenya, and provide recommendations for policy interventions that will lead to increased yield per acre. First, fertilizer subsidy programs should be implemented in regions where an additional unit of fertilizer applied to a maize field adds to incomes. For Kenya, such information is limited and the

state distribution of vouchers is not based on such indicators. Increased use of fertilizer in areas under moisture stress resulting from low rainfall or no irrigation infrastructure may not generate returns to cover the costs. The results from this study will provide policy makers with indicators that will assist in targeting assistance more effectively. We use a nationally representative panel household survey (1996/97, 1999/00, 2003/04, and 2006/07) and apply panel econometric methods that model for unobserved effects and mitigate biased estimates of model parameters and derived values that are associated with cross-sectional analyzes.

In this section we summarize the literature on some of the results from studies on increasing yields per acre in developing countries. Phimister and Roberts (2006) and Savadogo et al (1994) argue that raising productivity through a mix of input and infrastructure incentives will increase farm incomes and reduce poverty. Alene et al (2008) and Marennya and Barret (2009) use cross-sectional data to estimate a von Liebig quadratic-with-plateau yield response function for maize yield in Kenya and conclude that fertilizer application is not profitable even with subsidies. However, they are not able to model unobserved heterogeneity due to the cross-sectional nature of their data and therefore their results could be biased. Savadogo et al (1994) estimate the effect of animal traction and non-farm income on productivity in Burkina Faso modeling yield response as a quadratic function and find that the type of land preparation technology used has an effect on yields. Liu and Myers (2009) use a subset of the Kenya household panel data used in this study (fewer waves) to estimate technical efficiency and find that the estimates are not robust to functional form and that 10% of technical efficiency is accounted for by education, gender, distance to road, tenure, land size, and income. They also show that the parameters of the stochastic frontier are robust to different specification models of

the inefficiency as depicted in Alvarez et al (2006). However, they do not generate measures to guide fertilizer intensification efforts, a key objective of this study.

Abrar and Morrissey (2005) use Ethiopia data to show that excluding technical inefficiency when estimating response functions leads to inconsistent estimates. However, they assume profit maximization despite their data showing thin markets and poor integration and their profit model does not fit some zones. A number of papers conclude that specification issues like omission of relevant variables (Chhibber 1989) or oversimplifying production structures facing farmers by not capturing all variables (Ogbu and Gbetiouo 1990) result in biased results. Studies like those by Staatz, Dione, Dembele (1989), Weber et al (1988) find some differences in responses between larger and small farms.

1.1.2 Objectives

The principal objective of this essay is to provide estimates of maize fertilizer profitability for rural households, by region, across Kenya. Profitability measures include income added by the utilization of fertilizer. A second objective is to assess the estimated value of the marginal product of fertilizer, at observed application rates, relative to the cost of the fertilizer in order to assess whether additional income can be added by modifying fertilizer application rates. A third objective is to test particular features of the production relationships that are important in understanding farmer choices and for public policy such as the impact of the moisture stress environment on seed type choice and fertilizer productivity.

Profitability estimates and marginal products are based upon estimated production function for yield per hectare. The relationship between maize yields and nitrogen and phosphorus application rates, seed, labor, gender of the head of household, mixed versus single

crop plots, type of seed planted (hybrid and non-hybrid), different types of land preparation technology (tractor, ox, and manual), agro-zones, and soils is modeled. We also control for moisture stress with a variable that measures the proportion of twenty-day periods during the season with rainfall below forty millimeters³. This is very important indicator of the impact of availability of moisture and its effect on crop output.

We also test field anecdotes from some farmers that hybrid seed performs worse than non-hybrid or local seed under moisture stress. This was cited by some farmers as one reason why they have not switched completely to hybrid seed. They maintained that under stress conditions, local seed guarantees some harvest while hybrid seed may not. We will test for the significance of the interaction between moisture stress and type of maize seed planted to verify this claim.

1.1.3 Functional Form and Econometric Considerations

The data we use for estimation is described in section 1.3.1 and in more detail in the Appendix B (Table 2). The existence of household unobserved effects that may lead to biased estimates if not accounted for by the estimation procedure. Farmers are assumed to be price-takers in a competitive maize markets. The key assumption is that households make decisions based on unique soil quality conditions on their farms, expected prices, and managerial skills of decision-makers that are not observed in the sample information. These farmers face uncertain rainfall and price risks from shifting global and local demand and supply conditions. Such risks

³ Rainfall periods were broken down into 20-day periods in order to capture the importance of moisture throughout the growing period. Amounts below 40 mm of rainfall were taken as the threshold below which maize growth is constrained. These data was compiled from National Weather Service Climate Prediction Center (CPC) as a part of their Famine Early Warning System (FEWS) Project.

imply that farmers maximize a measure of utility based on expectations on prices and random factors such as rainfall.

It is important to control for the unobserved farm effects which can bias parameter estimates. To select the appropriate model, it is necessary to test for the presence of a relationship between explanatory variables and unobserved heterogeneity. A commonly used model assumes that unobserved effects (C_i) are constant across the panel time period but differ across observation units:

$$y_{it} = x_{it} \beta + C_i + u_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, N, \text{ where } x_{it} \text{ is } 1 \times k \text{ and contains}$$
 observable variables that change across t and interaction terms making the model flexible and β is $k \times 1$. Estimation hinges on whether the unobserved effects C_i are correlated with x_{it} or not. This problem is evident if $\text{cov}(x, C) \neq 0$ and C_i is part of the error structure, which makes pooled OLS biased and inconsistent.

For cross-section data, replacing the x_j that are correlated with C_i with a proxy variable, or using instrumental variables, is one way of approaching this problem. Under appropriate assumptions, the Correlated Random Effects (CRE) or fixed effects panel data approaches can generate consistent estimates in the presence of unobserved random variables or omitted variables. The Mundlak-Chamberlain approach (Mundlak 1978, Chamberlain 1982, 1984, Wooldridge 2002) controls for the relationship between independent variables and unobserved heterogeneity. The structure of the CRE model allows for a Wald test to select among suitable models.

The CRE replaces C_i with its projection onto the explanatory variables including the projection error term and assumes no particular distribution for $E(C_i | X_i)$. Unlike RE, this approach models a relationship between C_i and X_{it} as follows;

$C_i = \tau + \bar{X}_i \xi + a_i$, where $C_i \sim \text{Normal}(\tau + \bar{X}_i \xi, \sigma_a^2)$ and $\bar{X}_i$ is the average of X_{it} , $t = 1, \dots, T$, for each unit i.e. $\bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}$ for $i = 1 \dots N$. Also τ and ξ are constants, and σ_a^2 is the time-invariant variance for a_i assuming $a_i | X_i \sim \text{Normal}(0, \sigma_a^2)$ and $\text{cov}(u_{it}, a_i) = 0$ (Wooldridge 2002).

The CRE approach is useful in a number of ways. If the null $H_0: \xi = 0$ is rejected, then the estimation process favors CRE and FE over RE since the implication is that there is a relationship between the unobserved heterogeneity and the exogenous variables which, if not controlled for or eliminated, will result in inconsistent estimates. Under certain distributional assumptions, the CRE approach is preferred since it generates the same estimates for the time-varying variables as FE while also estimating the effect of time invariant variables, unlike FE; in this regard CRE nests FE. FE approach differences away time-constant variables and so does not provide these estimates. Both FE and CRE require sufficient variation in variables for identification of estimates (Table 1).

A decomposition of total variance for key variables into within and between components indicates significant within variation which means there is adequate within farm information to

estimate fixed and correlated random effects (CRE) models. They are relevant models for estimation of the parameters needed for meeting the objectives of the essay.

Table 1. Result of Variance Decomposition (between, within and total variations)

	N	P⁴	S	ST	G	MS	LP	CM
High Potential Region								
Between	73.52	18.97	4.75	0.09	0.43	0.02	0.60	0.04
Within	65.31	19.73	9.84	0.06	0.12	0.03	0.16	0.09
Low Potential Region								
Between	3.01	0.48	8.02	0.11	0.50	0.06	0.30	0.05
Within	1.95	0.45	11.49	0.11	0.20	0.07	0.13	0.09

Note: N=Nitrogen (kgs / acre) P=Phosphorous (kgs / acre) ST=Seed type (1=Hybrid, 0=traditional) G=Gender (1=Male, 2=Female) R=Rain (mm) MS=Moisture stress which indicates the fraction of 20 day periods with <40mm rain during the growing season LP=Land preparation method (1>manual 2=oxen 3=tractor) CM=Crop mix (1= mixed crop,2=mono crop plot).

Next we determine the form the regression model will take. Specification of the functional form is important since it influences the parameter estimates. We use a quadratic CRE model in levels (nominal). Though the quadratic logarithmic (translog) format offers appealing attributes, we do not use it for a number of reasons. The elasticities for a linear in logarithms regression can be extracted directly. The logarithmic model mitigates heteroscedasticity that frequently is associated with the yield-input rate relationship. Further, the estimated marginal products contain estimated unobserved farm effects; they do not for the quadratic in nominal variables model. However, zero rates of input, which occur with zero fertilizer rates for some households for one or more waves, are a challenge because the logarithm of zero is undefined. Thus approximations are required. Also, the rate of output at low input rates approaches zero; farms with zero rates of application for fertilizer have expected non-zero output

⁴ The $P_2 O_5$ content in fertilizer products was converted to P

(perhaps 40 to 50 percent of output at fertilizer rates that maximize expected yield) which also requires adjustments in the basic model by transforming observed yield.

We estimate a quadratic CRE model in nominal. Both fixed effects and random effects (RE) models are estimated to permit testing for exogeneity of the input rate variables. The results from the fixed effects model will be used in the estimation of marginal products (MP) and profitability measures if exogeneity is rejected.

1.1.3.1 Data and Variable Description

We use a four-wave panel household survey from rural Kenya collected through joint efforts of Egerton University's Tegemeo Institute and Michigan State University with funding from USAID. Table 2 gives a description of the variables that are used in this study.

Table 2. Description of Variables used in Production Function Regression

Variable	Description
Yield	Yield index (kg / acre)
N	Nitrogen nutrient content of fertilizers (kg / acre ⁵)
P	Phosphorous nutrient content of fertilizers (kg / acre)
L	Labor (days / acre): 1 labor day=8 hours of work by adult
S	Maize Seed (kg /acre)
R	Rainfall (mm) during the growing season
MS	Proportion of 20-day periods with <40mm rainfall during growing season
ST	Seed Type : 1=hybrid seed ; 2=non-hybrid seed
LP	Land Preparation: 1= manual hoe; 2=ox-plough; 3=tractor
CM	Crop Mix: 1=multiple crop plot; 2=single crop plot
G	Head of household: 1=male; 2=female with non-resident spouse
Soil	Five soil categories based on sand and clay content
Zone	Seven agro-ecological zones

⁵ We use acre instead of hectare (ha) because this is the common expression for size of land in Kenya. One hectare is equivalent to 2.47 acres

More information on the description and summary statistics of the data used for both Essay 1 and Essay 2 is available in Appendix 1 and Appendix 2 in Section A1.2.1, Table 6, Table 7, Table 10, and Table 12 to Table 17. The following section explains how the measure of yield was generated using relative crop prices as weights.

1.1.3.2 Generation of Yield Index for Mixed Plots

Approximately 85 percent of all plots planted with maize also contain at least one other crop i.e. two or more crops planted in the same plot. For instance, a maize-bean intercrop is a plot that contains either i) rows of maize and bean seeds planted in the same hole or ii) alternate rows of maize and beans. Inputs used on multi-crop plots are difficult to allocate to specific crops. In order to generate a measure of yield per acre for such plots, it is necessary to convert all output into a weighted index that represents all crops. Using relative prices as weights, all output in each plot is converted into maize-equivalent yield index as follows:

$$\sum_{i=1}^n q_{ijt} \frac{r_{it}}{P_{Mt}} \text{ where } q_{ijt} \text{ is the quantity of crop } i \text{ in plot } j, r_{it} \text{ is its price and } P_{Mt} \text{ is}$$

the price of maize output at time t , and n is the number of crops in the plot. This gives a weighted quantity of output for each plot which we then divide by the size of the plot to get “yield index” per acre (Liu and Myers, 2009).

The other challenge when dealing with mixed-crop plots is deciding on what “seed” measure to use. The quantity of maize seed planted per acre differs depending on intensity of inter-cropping and the type of seed planted (traditional or hybrid). The general recommendation is 10 kilograms of hybrid maize seed planted per acre. Since most farmers use manual means to plant seed and so the quantity planted per acre differ across farms, more so for traditional seed. This study uses maize seed for analysis and ignores planting material for the rest of the crops

planted on the same plot as maize. This is an area that needs further research to identify ways to capture the diversity and resource allocations for the mixed plots.

1.1.3.3 Empirical Model

The empirical model builds on a number of preliminary tests on a number of hypotheses. We test for the existence of different response functions resulting from the type of seed planted and agro-zones. We assume that the yield response to fertilizer uptake differs between hybrid and non-hybrid seed. The assertion by some farmers that non-hybrid seed withstands poor weather better than hybrid seed is also tested. If this latter is true then yield should be less responsive to hybrid seed compared to non-hybrid seed under poor moisture conditions. We also test for second order or indirect effect on yields from applying phosphate fertilizer alongside nitrogen as argued in Baanante (1997). This is the increase in yield due to the positive effect of phosphate fertilizer (P) on the response of crop yield to a unit of nitrogen input i.e.

$$\left(\frac{\partial(\partial \text{yield} / \partial N)}{\partial P} \right) \text{for N.}$$

Other hypotheses are based on responses related to whether plots are single or mixed crop. We hypothesize that the quantity of maize seed planted declines as the number of crops on the plot increases. Figure 6 was generated from the dataset and supports this argument indicating decreased quantity of maize seed as number of crops increase within the plot. Regression analysis will give us better insights to its significance.

$$\begin{aligned}
yield_{it} = & \beta_0 + \sum_{s=1}^6 (\beta_s x_{it} + \beta_{ss} x_{ist}^2 + \alpha_{ss} x_{ist} ST_{ist}) + \frac{1}{2} \sum_{s=1}^6 \sum_{j \neq s}^6 \beta_{sj} x_{ist} x_{ijt} \\
& + \sum_{m=1}^4 \beta_m z_{imt} + \beta_{nm} CM_{it} N_{it} + \beta_{sm} CM_{it} S_{it} + \sum_{h=1}^5 \beta_h Soil_h + \\
& \sum_{g=1}^3 \beta_g Zone_g + MC_i + u_{it}
\end{aligned}$$

The variables in the above model are described in Table 2 above and X denotes a set of continuous variables (N, P, S, L, R, MS) while Z stands for categorical variables (ST, CM, LP, G).

1.2.0 Results and Recommendations

The discussion of the results covers the yield response for continuous followed by categorical variables and then focusses on the role of nitrogen on increasing yields and incomes for households.

This section compares results from random effects estimation with and without Mundlak-Chamberlain (M-C) device for the high and low potential regions. Likelihood ratio and Wald tests reject the hypothesis of unobserved effects being independent of input rates which favors Correlated Random Effects (CRE) as the appropriate estimator. Table 4 in the Appendix contains the regression results from a quadratic fit from which the marginal effects in Table 3 are derived. Table 2 provides a full description of the abbreviated variable names in Table 3 and Table 4.

As expected the marginal productivity for the nitrogen (N), phosphorus (P), labor (L), seed (S), and moisture stress have the expected signs. The effects of total annual rainfall and moisture stress in the high potential region have the expected signs but moisture stress is not statistically significantly different from zero. However, both rainfall and moisture stress variables are negative and significant for the low potential region (Table 3 and Table 4); the highly positive collinear relationship between the rainfall and moisture stress measures impacts the precision of the estimates (Table 8 and Table 9). Moisture stress is a more serious problem in the low potential region where climatic conditions are more volatile.

Table 3. Marginal Products of Inputs Averaged across Farms⁶ from Correlated RE and RE Regressions for Continuous Variables (yield index as dependent variable)

VARIABLES	High Potential Region		Low Potential Region	
	CRE	RE	CRE	RE
N	18.15***	21.39***	21.94	17.31
	(2.52)	(2.22)	(27.56)	(26.72)
P	2.30	-2.54	-64.34	-48.39
	(4.24)	(3.74)	(53.95)	(52.76)
S	42.85***	45.66***	42.39***	42.17***
	(4.22)	(3.92)	(3.73)	(3.42)
L	2.77***	2.57***	4.07***	3.79***
	(0.45)	(0.43)	(0.47)	(0.45)
R	0.30***	0.20**	-0.33**	-0.29*
	(0.10)	(0.09)	(0.16)	(0.15)
MS	6.17	-31.75	-286.30***	-273.74***
	(101.30)	(98.38)	(84.73)	(78.71)
Observations	3,127	3,127	1,371	1,371
Number of households	828	828	363	363

Note: Standard errors are in parentheses: *** p<0.01, ** p<0.05, * p<0.1. CRE=Correlated Random Effects; RE=Random Effects.

Table 2 explains the notation for variables used here. The categorical (dummies for gender of head of household, maize seed type, mixed crop, soil and AEZs) variable effects are discussed in the following sections.

The marginal contribution of a kilogram of seed is forty kilograms of output (Table 3). It is difficult to interpret the influence of increasing quantity of maize seed (S) planted per acre due to complications from mixed cropping. Approximately 85% of all plots surveyed contained maize planted on same plot with other crops (Figure 6). Quantity of seed is a “nuisance” variable whose effects are difficult to interpret conclusively and this might be an area for future research.

⁶ Estimated using the “margins” command in STATA 12 Statistical Software.

1.2.1 Effect of Nitrogen Application Rates on Yields and Household income

In the following sections we look at the implication of the estimation results for nitrogen input on yields and household incomes. The analysis looks at both the marginal and average contribution to income by Nitrogen (N) at the household level. The marginal products are estimated for an additional unit of nitrogen fertilizer on yields.

Nitrogen (N) has a significant positive effect on yield index for the high potential region as shown by regression results in Table 3. This is positive but not significant for the low potential region which has very low fertilizer adoption rates relative to the high potential region. The interaction of N and P has a positive effect on yields which is an indication of some synergy resulting from using both nitrogen and phosphate fertilizers in Kenya.

Figure 5 shows the fit of yield and nitrogen nutrient for both high and low potential regions which depict this relationship while holding all other inputs at their means. The graph gives the predicted regression line including a 95% confidence interval. The production function for the Low Potential region reflects the low rainfall and poor soil conditions which justifies the choice to split the data into two groups for analysis. It is also clear that using additional nitrogen fertilizer contributes to increased production in both regions. The sections below will study whether it is profitable to use nitrogen fertilizer in these areas.

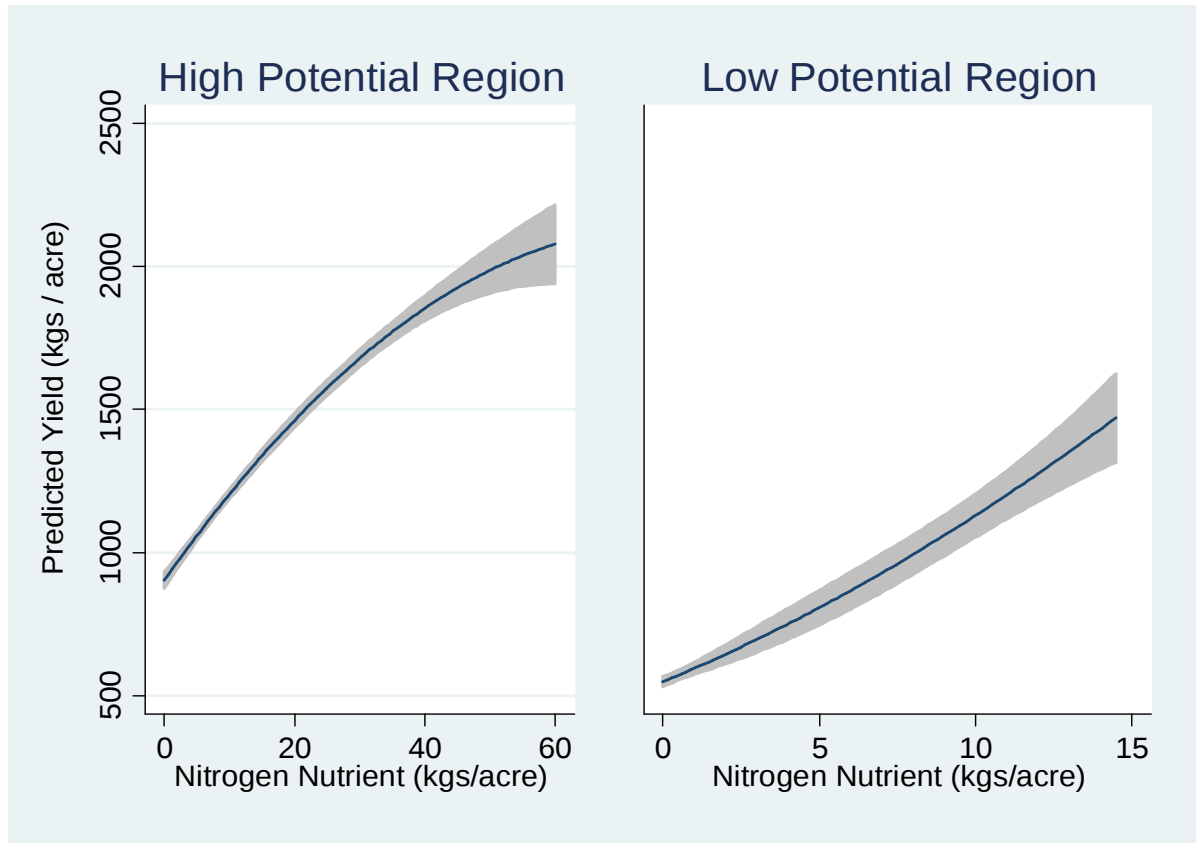
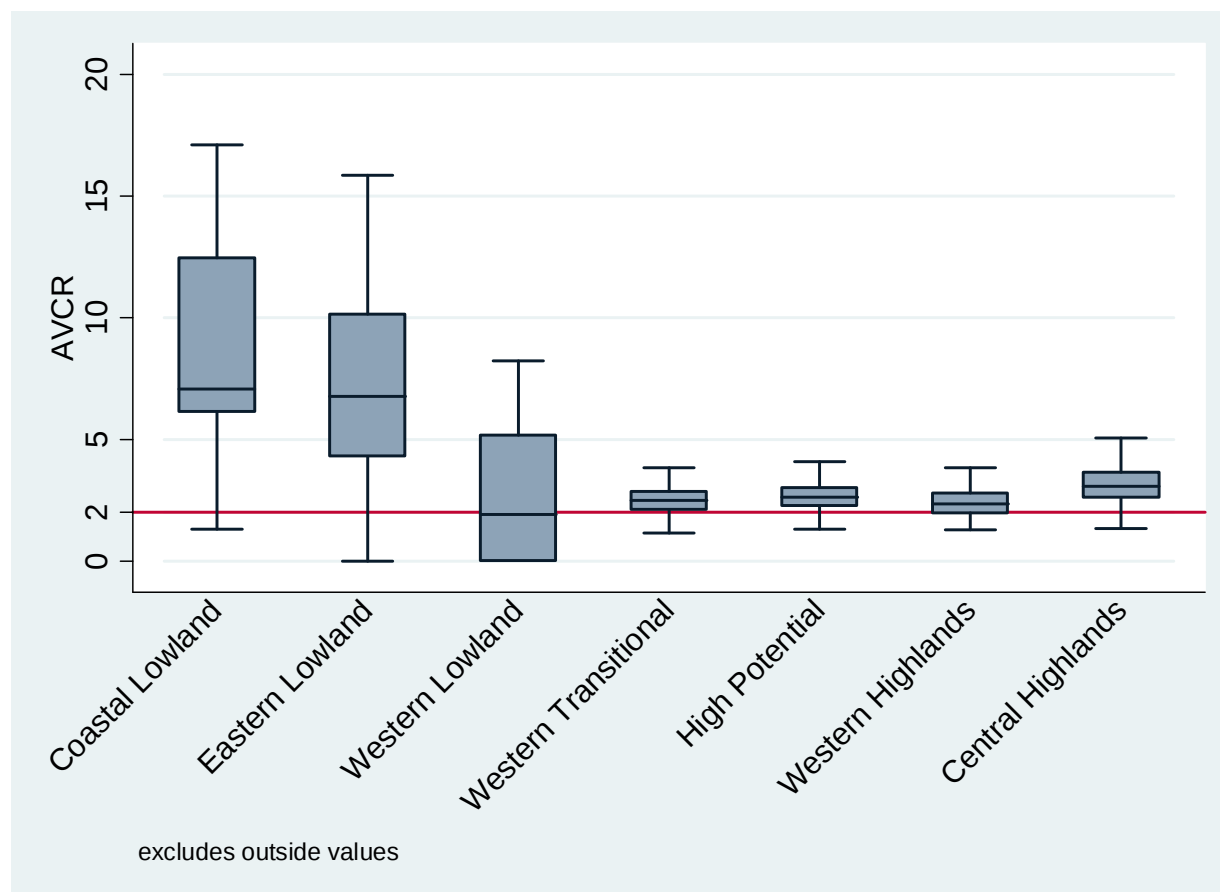


Figure 5. Predicted Yield versus Nitrogen

We test the hypothesis on the second order effects for phosphorus when interacting with nitrogen and find a positive but statistically insignificant relationship, implying a possible increase in yield due to the positive effect of phosphate fertilizer (P) on the response of maize yield to a unit of nitrogen nutrient (Baanante 1997). However, the high collinearity between nitrogen and phosphate fertilizers can influence the precision of these estimates.

We also generate the marginal and average products of N in order to derive household value cost ratios and income attributable to N (Table 5 in the Appendix and Figure 6 and Figure 7 below). These ratios are measures of the relationship between unit values and costs. The MVCR is the ratio of the value of marginal product of N to the cost of an additional unit of N i.e.

$MP_N * \frac{P_M}{P_N}$ where MP_N , P_M and P_N are respectively marginal product of N, maize price per kg and N price per kg. The AVCR is generated in a similar way using the “average product” (AP) at household level and then using available prices and generating averages at district levels.



Note: Districts within agro-zones are : Coastal Lowlands (Kilifi, Kwale), Eastern Lowlands (Taita Taveta, Kitui, Machakos, Makueni, Mwingi), Western Lowlands (Kisumu, Siaya), Western Transitional (Bungoma, Kakamega), High Potential (Bungoma, Kakamega, Bomet, Nakuru, Narok, Trans Nzoia, Uasin Gishu), Western Highlands (Kisii, Vihiga), Central Highlands (Meru, Muranga, Nyeri). This graph was generated using the “margins” command in STATA 12 Software.

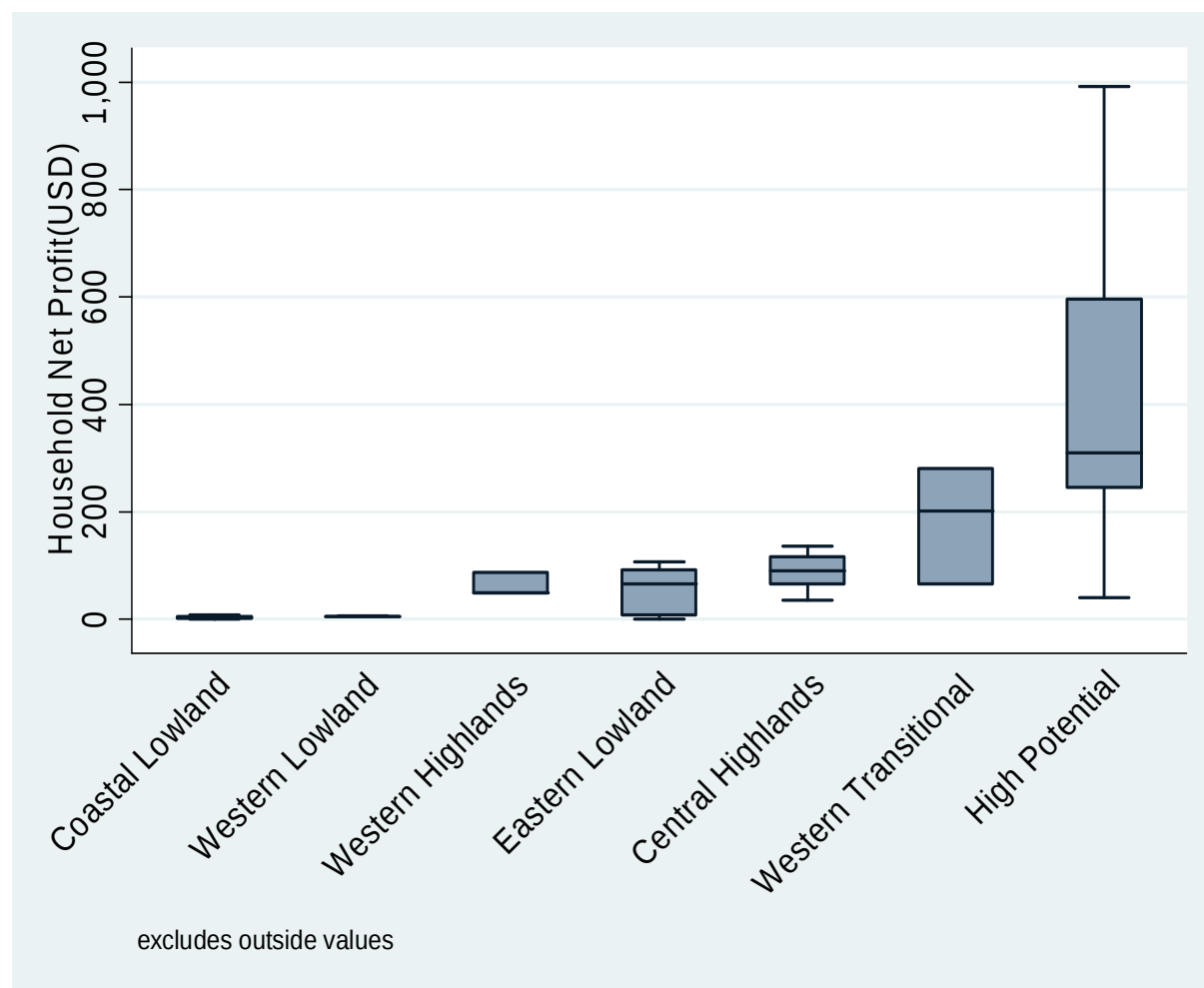
Figure 6. Marginal Value Cost Ratio for Nitrogen by Agro-Zone

Yanggen et al (1998) argues that a minimum AVCR of one is necessary but not sufficient for farmers to consider using fertilizer and a AVCR of greater than two may be required if there are significant risks involved. Marginal cost ratios are lowest in some of the districts that produce surplus maize sold in markets (Trans Nzoia and Uasin Gishu) probably because they are close to their optimal application rates. These are the districts with the highest nitrogen application rates per acre in the sample (High Potential Zone). The lowlands (coastal, eastern, and western) are classified as low potential due to their poor climatic conditions and low improved technology adoption compared to other zones. Except for the Western Lowland zone, the remaining zones meet the threshold for profitability indicated by a horizontal line in Figure 6.

The horizontal line in the box is the median while the lower and upper ends of the box represent the inter-quartile range (IQR), a measure of dispersion. The low potential zones (coastal and eastern) show relatively more dispersion. Interestingly, the low potential regions have relatively higher values for MVCR implying the potential for fertilizer use exists but empirical evidence reveals low percent of farmers adopting fertilizer and low application rates per acre. This is not happening for some households due to constraints that are dealt with in more detail in Essay 2 including production risks due to high moisture stress. Essay 2 estimates the effect of various variables on the demand for fertilizer, both on the decision to buy and how much to apply per acre. Distance to fertilizer seller, relative prices, distance to extension service and good roads, wealth, and education levels are some of the factors that explain fertilizer demand differences across the sample units.

Household income attributable to nitrogen fertilizer is generated from the difference between predicted yields with and without fertilizer at the household level. This difference is divided by the quantity of nitrogen fertilizer to get the “average product” (AP) of N. In its

simplest form, if the value of the AP is greater than the average cost of N, then one should use more N. The incomes are estimated as a product of the value of average product (VAP) minus the cost of a unit of fertilizer multiplied by the quantity of fertilizer used by the household.



Note: Districts within agro-zones are : Coastal Lowlands (Kilifi, Kwale), Eastern Lowlands (Taita Taveta, Kitui, Machakos, Makueni, Mwingi), Western Lowlands (Kisumu, Siaya), Western Transitional (Bungoma, Kakamega), High Potential (Bungoma, Kakamega, Bomet, Nakuru, Narok, Trans Nzoia, Uasin Gishu), Western Highlands (Kisii, Vihiga), Central Highlands (Meru, Muranga, Nyeri). This graph was generated using the “margins” command in STATA 12 Software.

Figure 7. Household Net Profit Added from Nitrogen Use (USD)

Most districts generate positive income attributed to fertilizer use on maize fields. The total household average profits are highest for the high potential zones since they consume more fertilizer (Figure 7). The highest incomes attributable to fertilizer use were in Trans Nzoia, Uasin Gishua, and Nakuru (the “bread basket” of the country where most surplus maize is produced), and Bomet, Nyeri, Kakamega. Within the low potential areas, incomes were higher for Machakos and Makueni (Eastern Lowlands) but zero for Kisumu and Siaya (Western Lowlands Zone) and Coastal Lowlands. This region has very low adoption and application rates for fertilizers. As mentioned above, apart from the unfavorable climatic conditions, other constraints exist that impede fertilizer use (these are explained extensively in Essay 2).

For a long time the agriculture extension system in Kenya has recommended one 50-kg bag of basal or planting fertilizer (usually DAP) and one 50-kg bag of top dressing (usually CAN) per acre which translates to approximately 23 kilograms of N per acre. The blanket recommendations by CIMMYT in 1994 were 37 kilograms of N per acre (Jewell et al 1994). It is important to note that there are a number of private fertilizer dealers that are currently offering soil-testing services and fertilizers that are tailored to specific soils conditions and crops. The Kenya recommendations for nitrogen are barely met in only two of districts that were studied, Trans Nzoia and Uasin Gishu. These are some of the few districts that account for most of the maize that gets to the market in Kenya (the others are Nakuru, Narok, and Bomet). The average nitrogen nutrient application rate per acre is 9 kilograms for high potential region (Table 7).

1.2.2 Effect of Gender on Yields

From the regression, the yield index in the high potential region is lower (8%) for households headed by females compared to those headed by males as shown in Figure below. The difference is statistically significantly different from zero.

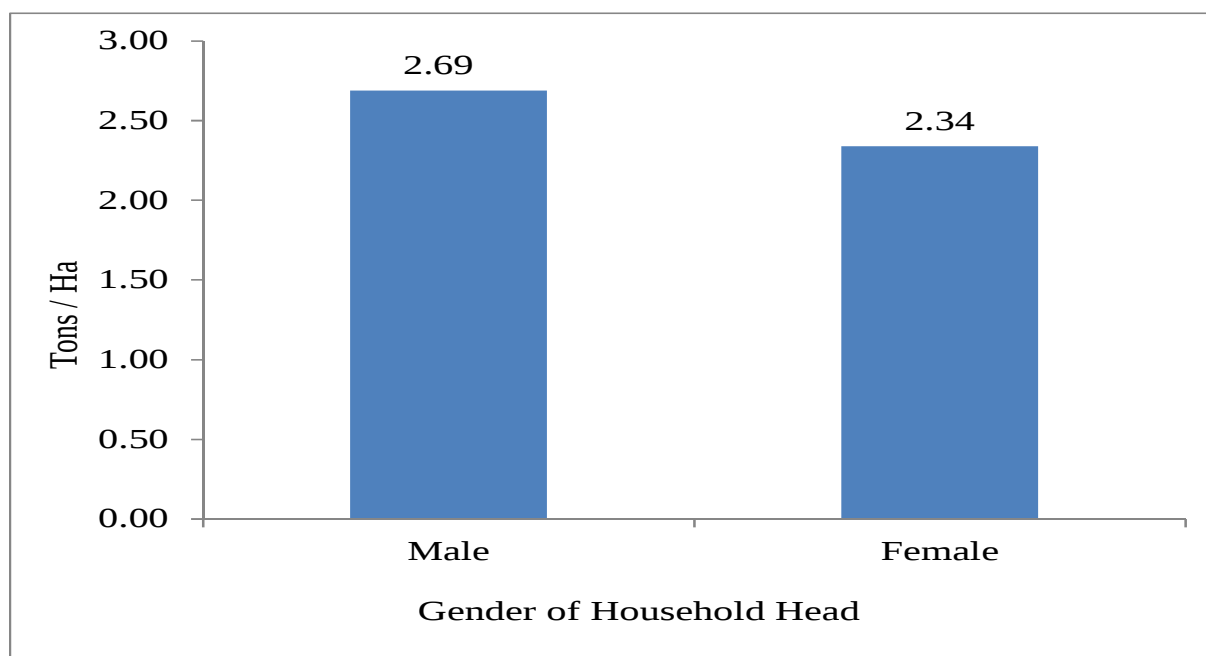


Figure 8. Maize Yield by Gender of Head of Household for the High Potential Region

Female-headed households without a spouse in the low potential region have significantly lower yields than male and female-headed households with a spouse by approximately 70 kilos per acre. This may reflect particular constraints associated with single-female-head households which in some cases are faced with relatively fewer resources like land and other assets.

1.2.3 Effect of Agro-Ecological Zones on Yields

There are significant differences across AEZs with the High Potential zone having the highest yield index per hectare. As expected, the high potential Maize (HPM) zone has higher yields (0.45 tons per hectare) than the Western Transitional region which in turn has higher yields than the Central and Western Highlands (Figure 9). For the low potential region, there is no significant difference in yields across zones.

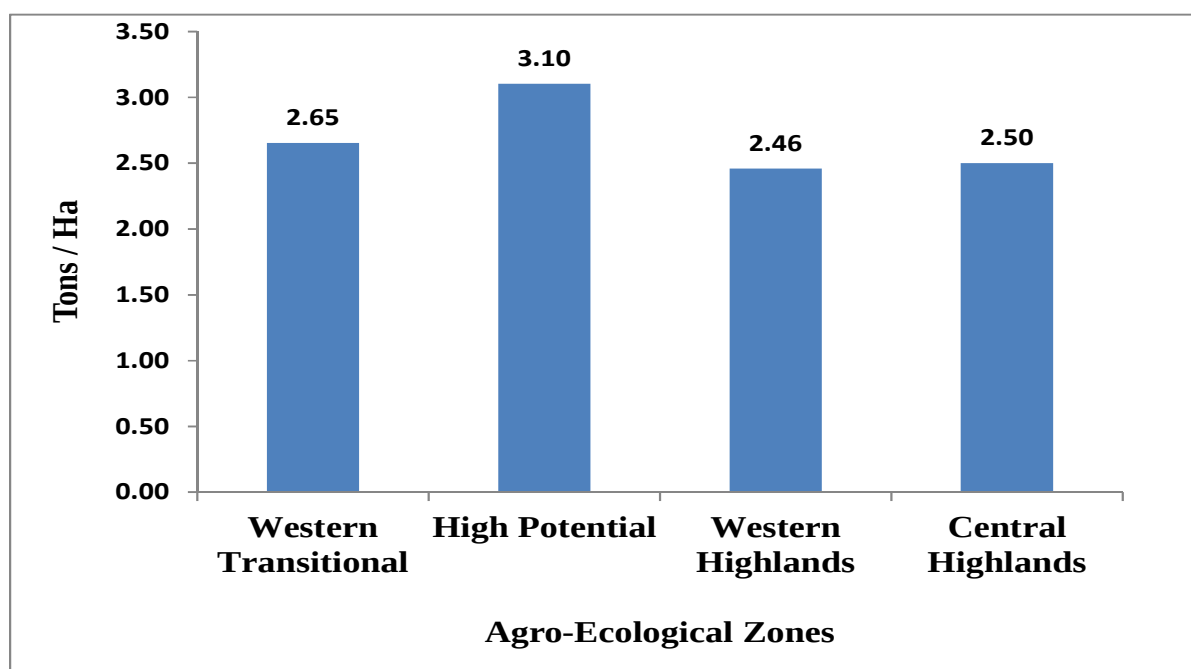


Figure 9. Maize Yield by Agro-Ecological Zone within the High Potential Area

There is a relationship amongst exogenous variables AEZs, soils and rainfall data used in the model. This might imply that though estimates for these variables are unbiased their relative and joint effects may not be reliable. Yields are not statistically significantly different across soil types for low potential region. For high potential region, the Regosol and Ranker types of soils

have higher yields compared to Cambisols and Phaeozem types⁷. It is important to note that the soil profile information used for this study is based on an old study done in the 1970s covering large areas of land. There are on-going current efforts to update soil information throughout the country.

1.2.4 Effect of Seed on yield Index

There are differences in yield depending on type of seed that is planted (hybrid or non-hybrid seed). Yield index per acre in the High Potential area are on average 15 percent higher with hybrid compared to non-hybrid seed. This is a significant difference of approximately 0.45 tons for hybrid compared to traditional seed users. Though this gap between hybrid and traditional seed is higher for Low Potential areas, the index is affected relatively more negatively by moisture stress possibly from their highly positive collinear relationship (Table 8 and Table 9) implying risks to fertilizer use in this drier region. This is reflected in low fertilizer adoption rates in this region. Availability of risk-mitigating factors (irrigation or drought tolerant crop varieties) may provide opportunities for increased fertilizer use in this drier area.

In the Low Potential region there is a significant and positive effect from interaction of nitrogen and hybrid seed resulting in 60 kilograms of output compared to traditional seed; this is positive but insignificant in the high potential region. This latter can be attributed to the relatively high adoption of hybrid seed (over 95 percent of households use hybrid seed) which offers little variability in the sample.

⁷ Soil classifications are explained under Section A1.2.1 in Appendix A and Table 10.

The yield index response for hybrid seed under moisture stress is negative for all regions (Table 4). It is, however, statistically significantly negative for high potential region (Table 4). These findings relate to the assertion by farmers that local seed does relatively better under moisture stress than hybrid seed.

However, it is more difficult to interpret the influence of increasing seeding rate due to complications from mixed cropping, where maize is grown in the same field with other crops. The effect on yield from increasing the quantity of seed is significantly positive (

Table 3) and declining as seeding rate is increased (Table 4). The positive effect on yield from increasing quantity of maize seed planted could imply that raising seed quantities will increase production. However, there is a threshold quantity above which yields will decline due to poor agronomic conditions; it is recommended that farmers apply 10 kilograms of maize per acre (sample rates are 9.5 and 7.6 kilograms per acre respectively for high and low potential regions). As expected, seeding rate per acre is less for mixed crop compared to single crop plots. Yields on plots planted with a single crop have lower response to N than mixed crop plots (13 percent lower) indicating higher N utilization efficiency on mixed cropped plots for households in the high potential region⁸. This implies that MP, given a rate of N, is higher on monocrop plots. Seeding rate is a “nuisance” variable whose effects are difficult to interpret conclusively and this might be an area for future research.

⁸ Note that as explained in the Appendix, yield is an index of output from all crops in the plot including maize.

1.3.0 Conclusions and Recommendations

The results reveal a number of areas requiring attention in order to raise maize productivity and profitability. Use of nitrogen has positive effect on incomes for most households. Though the MVCR differ substantially across districts, for most of these areas, these values are greater than 2, suggesting a basis for additional increases in nitrogen application rates

Essay 2 provides insights into why despite favorable MVCR and AVCR farmers in some of these areas may not be using fertilizer. Some of these constraints include: distance to nearest fertilizer seller, moisture stress, level of wealth, education, and relative prices. These explain why districts in drier areas have some of the highest “profitability” indicators (MVCR, AVCR) but adoption rates are relatively low compared to other districts.

Joint use of nitrogen fertilizer and hybrid seed has positive effect on yields. An integrated approach with holistic view of farm incentives should include the appropriate combination of input technologies for improved performance in yields. Therefore, encouraging increased fertilizer use should go together with adoption of technologies that, taken together, increase yields.

Yields are higher for plots planted with hybrid than non-hybrid seed by 17 percent. In areas where moisture stress is not a constraint, hybrid seed adoption is relatively higher than other areas. The results indicate that non-hybrid seed does better under moisture stress relative to hybrid seed. It is therefore important that farmers in drier areas get access to seeds that are tolerant to moisture stress. This needs dissemination of information by extension workers and policy intervention to encourage access to appropriate technologies.

There is variation in yield across agro-zones, soil types, and gender of the head of household. For both high and low potential regions, households headed by males have higher yields than those headed by females. It is interesting that female-headed households without a spouse have lower yields than female or male-headed households with a spouse. This is an area that requires further research to understand the reasons for this disparity.

APPENDIX

APPENDIX 1

A1.1: Regression Results and Analysis

Table 4. Results from CRE and RE Regressions for High and Low Potential Regions

	High Potential Region		Low Potential Region	
VARIABLES	CRE	RE	CRE	RE
N	49.30***	51.87***	150.27***	147.67***
	(9.59)	(9.48)	(46.72)	(45.77)
P	-56.80***	-60.11***	-214.87***	-199.29**
	(18.24)	(17.95)	(79.58)	(79.03)
S	48.19**	57.59***	76.69***	77.37***
	(21.38)	(21.19)	(10.71)	(10.52)
L	3.95*	4.43**	5.34***	4.99***
	(2.10)	(2.09)	(1.25)	(1.25)
R	-1.71***	-1.63***	0.26	0.26
	(0.53)	(0.52)	(0.52)	(0.49)
MS	-1,915.56***	-1,724.84***	-251.59	-230.08
	(434.16)	(429.66)	(250.97)	(241.84)
N#N	-0.16	-0.19*	-7.98**	-7.89**
	(0.10)	(0.10)	(3.50)	(3.50)
P#P	0.47	0.51	21.43	21.66
	(0.43)	(0.42)	(20.58)	(20.53)
S#S	-3.11***	-3.15***	-1.61***	-1.71***
	(0.65)	(0.65)	(0.51)	(0.50)
L#L	-0.01	-0.00	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
R#R	0.00	0.00	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
MS#MS	754.49**	637.24*	-534.07**	-516.55**
	(333.95)	(331.57)	(230.37)	(217.79)
N#P	0.18	0.24	11.54	11.66
	(0.29)	(0.29)	(12.16)	(12.13)

Table 4 (cont'd)

N#S	-0.03	-0.07	-8.17**	-8.74**
	(0.46)	(0.46)	(3.41)	(3.40)
N#L	-0.11***	-0.11***	-0.62*	-0.61
	(0.04)	(0.04)	(0.37)	(0.37)
N#R	0.00	0.00	-0.10	-0.10
	(0.01)	(0.01)	(0.14)	(0.14)
N#MS	-22.56***	-21.64***	-35.62	-34.56
	(6.86)	(6.85)	(54.22)	(54.00)
P#S	1.00	0.96	6.63	7.26
	(0.86)	(0.87)	(6.12)	(6.10)
P#L	-0.07	-0.07	0.98	0.96
	(0.08)	(0.08)	(0.78)	(0.78)
P#R	0.04***	0.04***	0.05	0.04
	(0.01)	(0.01)	(0.18)	(0.18)
P#MS	21.93*	22.25*	83.64	85.43
	(13.12)	(13.11)	(102.05)	(101.92)
S#L	0.27***	0.25**	0.11	0.13*
	(0.10)	(0.10)	(0.08)	(0.08)
S#R	0.01	0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
S#MS	32.13**	27.83*	-20.16**	-20.15**
	(15.48)	(15.45)	(10.06)	(10.06)
L#R	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
L#MS	-4.08***	-4.63***	-0.66	-0.59
	(1.57)	(1.57)	(1.41)	(1.41)
R#MS	1.48***	1.50***	2.36***	2.27***
	(0.43)	(0.42)	(0.45)	(0.41)
ST	259.25	259.38	1.27	9.02
	(172.81)	(172.82)	(60.33)	(59.95)
CM	-140.26	-146.29	-134.65**	-145.24**
	(89.12)	(89.14)	(60.98)	(60.93)
LP:				
Ox	-58.91	-82.83**	9.70	38.77
	(39.45)	(38.96)	(27.67)	(24.72)

Table 4 (cont'd)

Tractor	57.23	60.48	-0.76	19.19
	(39.49)	(38.91)	(41.79)	(40.33)
Gender of Head:				
Female, with Spouse	-84.72**	-81.62**	-36.93	-33.74
	(38.89)	(38.91)	(29.79)	(29.72)
Female, without Spouse	-7.81	-12.30	-71.65***	-71.67***
	(32.92)	(32.94)	(27.10)	(27.01)
Soils:				
phaeozems	-155.68***	-157.61***	-12.29	53.94
	(39.25)	(38.94)	(63.56)	(55.91)
Cambisols	-201.99***	-255.85***	-24.46	-39.73
	(55.73)	(53.59)	(79.97)	(75.46)
Regosols	8.05	-34.34	0.00	0.00
	(45.48)	(43.46)	(0.00)	(0.00)
Others	-22.93	-24.30	-110.30**	-76.24**
	(65.00)	(64.29)	(45.64)	(34.24)
AEZs:				
Coastal Lowland			150.14	-148.60***
			(141.69)	(49.18)
Eastern Lowland			16.01	-327.64***
			(164.96)	(47.61)
Western Transitional	-182.27***	-226.32***		
	(56.26)	(47.98)		
Western Highlands	-261.30***	-324.41***		
	(64.73)	(52.71)		
Central Highlands	-244.00***	-175.40***		
	(63.15)	(54.84)		
ST#N	-8.09	-8.22	37.04	34.81
	(5.32)	(5.32)	(31.40)	(31.38)
ST#P	3.94	3.02	-47.34	-43.45
	(9.60)	(9.59)	(54.51)	(54.47)
ST#S	4.05	5.18	19.87***	19.24***
	(8.91)	(8.93)	(6.68)	(6.66)
ST#L	2.74***	2.76***	0.58	0.54
	(0.96)	(0.96)	(0.63)	(0.63)

Table 4 (cont'd)

ST#R	-0.07	-0.07	0.11	0.14
	(0.16)	(0.16)	(0.13)	(0.13)
ST#MS	-117.79	-129.48	-51.65	-70.50
	(145.44)	(145.54)	(81.76)	(81.29)
CM#N	-4.59**	-4.57**	-3.75	-1.86
	(2.29)	(2.30)	(14.28)	(14.26)
CM#S	2.12	2.64	-11.67	-10.66
	(9.06)	(9.07)	(7.24)	(7.24)
M-C:				
N	7.36**		-11.58	
	(2.93)		(20.27)	
P	-9.61*		59.05	
	(5.75)		(45.33)	
S	10.64		-3.53	
	(7.57)		(5.41)	
L	0.94		-1.10	
	(0.81)		(0.70)	
R	-0.75**		0.86*	
	(0.32)		(0.51)	
MS	116.16		120.66	
	(303.09)		(301.42)	
Constant	1,651.27***	1,419.60***	10.67	478.78***
	(303.32)	(270.15)	(257.30)	(80.01)
Observations	3,127	3,127	1,371	1,371
Number of households	828	828	363	363

Standard errors are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. CRE=Correlated Random Effects; RE=Random Effects. Table 1 and Table 6 (below) explain the notation for variables used here. The symbol “#” denotes interaction between two variables. Note that for categorical variables, one category is dropped and the remaining are compared to the dropped category. Female dummies: dummy for “1=male headed household” is dropped so this is the base for comparison for the other two categories. CM=mixed crop category dropped, ST=non-hybrid seed category dropped, Zone dummies: dropped high potential zone (6) for the high potential region and western lowland zone (3) for the low potential region; dropped soil type 4 for both zones. To avoid clutter the soils are explained in Table 10 and agro-ecological zones in section A1.2.1 M-C stands for Mundlak-Chamerlain device for unobserved heterogeneity.

Table 5. Marginal and Average Cost Ratios and Incomes

Region / District	MVCR	AVCR	Household Income Attributed to Nitrogen (US\$) [(VAP-Cost)*Qn]
Low Potential:			
Makueni	4.7	2.4	9.4
Mwingi	4.8	3.9	1.7
Kilifi	7.4	5.3	0.8
Machakos	4.0	3.5	6.4
Kisumu	0.0	0.3	0.0
High Potential:			
Narok	2.5	2.7	2.4
Meru	2.5	1.1	5.8
Bomet	2.3	1.7	15.4
Nyeri	2.8	1.6	18.2
Nakuru	2.4	2.3	20.6
Kisii	1.7	1.0	1.9
Muranga	2.7	1.1	3.2
Uasin Gishu	1.8	2.0	36.4
Vihiga	2.0	1.3	4.3
Kakamega	1.9	1.7	18.2
Bungoma	1.6	1.6	21.3
Trans Nzoia	1.7	1.7	27.6

Note: Qn is the average household quantity of nitrogen, VAP=Value of average product, cost=cost of a kilogram of nitrogen

A1.2: Additional Data Analysis

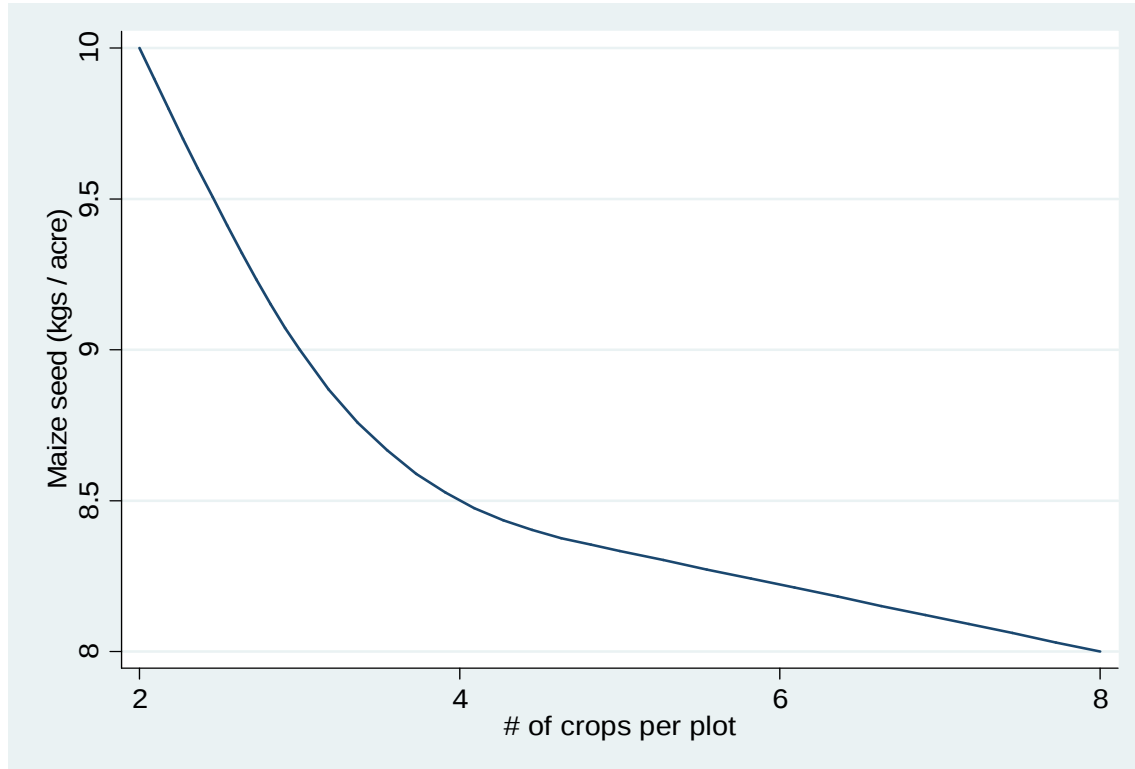


Figure 10. Quantity of maize seed as number of crops planted in plot increases

Table 6. Description of Variables used in Production Function Regression

Variable	Description*
Yield	Yield index (kg / acre)
N	Nitrogen nutrient content of fertilizers (kg / acre)
P	Phosphorous nutrient content of fertilizers (kg / acre)
L	Labor (days / acre): 1 labor day=8 hours of work by adult
S	Maize Seed (kg /acre)
R	Rainfall (mm) during the growing season

Table 6 (cont'd)

MS	Proportion of 20-day periods with <40mm rainfall during growing season
ST	Seed Type Dummy: 1=hybrid seed ; 2=non-hybrid seed
LP	Land Preparation Dummy: 1= manual hoe; 2=ox-plough; 3=tractor
CM	Crop Mix Dummy: 1=multiple crop plot; 2=single crop plot
G	1=male head of household 2=female with non-resident spouse
Soil	5 types of soil categories based on sand and clay content
Zone	7 different agro-ecological zones

Table 7. Means of Variables used in Production Function (1997-2007)

	1997	2000	2004	2007	Total
Yield	973	1196	1275	1442	1233
N	9.4	12.3	13.0	13.8	12.2
P	6.1	7.3	7.23	7.63	7.12
L	31.7	--	56.6	48.3	35.4
S	9.8	9.5	9.5	9.8	9.6
R	316	637	564	535	522
MS	0.00	0.21	0.36	0.24	0.22
ST:					
Hybrid (1)	0.84	0.84	0.79	0.83	0.82
Non-Hybrid (2)	0.16	0.16	0.21	0.17	0.18
LP:					
Manual (1)	0.50	0.52	0.50	0.50	0.50
Ox-Plough (2)	0.24	0.25	0.26	0.26	0.26
Tractor (3)	0.26	0.23	0.24	0.24	0.24
CM:					
Mixed Crop Plots (1)	0.79	0.88	0.86	0.82	0.84
Single crop (2)	0.21	0.12	0.14	0.18	0.16

Note: Labor data not collected in 2000. The numbers in parentheses next to the categorical variable classifications denote numerical values labels assigned to each category and used in regression results in Table 4.

Table 8. High Potential Region: Correlation Matrix for Continuous Variables

	Yield	N	P	L	S	R	MS
Yield	1.00						
N	0.36*	1.00					
P	0.31*	0.72*	1.00				
L	0.17*	0.05*	0.06*	1.00			
S	0.32*	0.18*	0.19*	0.10*	1.00		
R	-0.01	0.07*	0.08*	-0.06*	-0.04*	1.00	
MS	-0.20*	-0.04	0.00	-0.16*	0.01	0.23*	1.00

Note: significance is indicated at 5%. These variables are described in Table 1 and Table 6. Nitrogen (N), Phosphate (P), Labor (L), Seed (S), Rainfall (R), and Moisture Stress (MS). The table does not include categorical variables like agro-zones, soil types, seed type (ST), crop mix (CM), Land Preparation (LP), or gender of household head (G) to avoid clutter.

Table 9. Low Potential Region: Correlation Matrix for Continuous Variables

	Yield	N	P	L	S	R	MS
Yield	1.00						
N	0.25*	1.00					
P	0.21*	0.83*	1.00				
L	0.07*	0.03	0.03	1.00			
S	0.32*	0.09*	0.08*	0.04	1.00		
R	-0.08	-0.23*	-0.18*	0.20*	0.13*	1.00	
MS	-0.21*	-0.28*	-0.22*	0.16*	0.06*	0.65*	1.00

Note: significance is indicated at 5%. These variables are described in Table 1 and Table 6. Nitrogen (N), Phosphate (P), Labor (L), Seed (S), Rainfall (R), and Moisture Stress (MS). The table does not include categorical variables like agro-zones, soil types, seed type (ST), crop mix (CM), Land Preparation (LP), or gender of household head (G) to avoid clutter.

A1.2.1 Agro-zones, Rainfall, and Soil Data

It is important to capture effects of environmental factors to avoid endogeneity and therefore biased estimates if these are left out. Farmers are assumed to make resource decisions based on their environment (soils, rainfall, and other agro-ecological conditions) and these input decisions will affect yields. If these factors are not included, they will be in the error term and

therefore correlated with the explanatory variables. Following the theory of expectations, we use average of rainfall for all previous seasons including the current period for Essay 1 (similarly for the moisture stress variable which is derived from rainfall). For instance for the year 1997 we use average of 1995 to 1997, for the year 2000 we use rainfall from 1995 to 2000, for the year 2004, we use data from 1995 to 2004 and so on since rainfall data runs as far back as 1995. . However, for Essay 2 we used lagged rainfall; therefore for 1997 we use year 1996, for 2000 we use 1999, etc.

The soil file contains a number of variables including type of soil, soil description, landform, drainage, depth, geology, % silt, % clay, % sand with identifiers at village, division, district, zone, and province level, a careful analysis reveals that these variables vary more at district level (not village level as indicated). There is no sufficient variation in these variables to justify classification at levels below the district. The variables soil type⁹, soil description, landform, drainage, depth, and geology are definitely very collinear, changing only across districts; therefore using one of them is sufficient. Drainage is less variable than depth for example. Some two villages in Kisumu and Trans Nzoia districts have no soil data. We do not use district dummies alongside soil variables to avoid near perfect collinearity.

Soil types were re-grouped into 5 clusters by combining soils that had been classified under different names but which have similar characteristics (Table 13). The classifications that were used in this dataset contained similar soil types under different classifications due to differences in soil names between American and European-based soil survey organizations. Kenya soil map was produced by Kenya Soil Survey and the Ministry of Agriculture Nairobi

⁹ This study benefited from the experience and knowledge of GIS experts, agronomists, and soil scientists at both Michigan State University (MSU) and the International Center for Soil Fertility and Agricultural Development (IFDC). This effort culminated in re-classification of the soil types into five categories.

from topography work done by Defense Mapping Agency, Aero Space Center, St. Louis AFS, Missouri and scans by Canada Land Data Systems Division, Land Directorate, Department of Environment, Ottawa, Canada

Table 10. Depth and Drainage Characteristics of Soil Types

New Groups	Soils Types	Mean Depth	Mean Drainage
1	Phaeozems and Luvisols	2	9
2	Cambisols, Ferralsols, Vertisols	4	3
3	Regosols	4	8
4	Rankers	3	8
5	Greyzem, podzols, and Solonetz	3	4

Note: Extent of drainage ranges from 1=very poorly drained to 11=excessively drained. Depth ranges from 1=shallow to 8=extremely deep.

Total rainfall amounts for the season were extrapolated to villages from nearest weather stations and satellite imagery based on household and village GPS coordinates¹⁰. This variable has little variation within villages and more across villages and so is treated as a village level variable. Note that the variation across villages emanates more from satellite imagery differences rather than diversity in weather stations (these are few in Kenya). We generated a moving average of total rainfall during the growing season at village level from all previous years excluding the current year to represent long-run rainfall patterns. An additional variable was generated that gave the fraction of 20-day periods during the maize growing period with rainfall

¹⁰ An additional set of rainfall data based on GPS coordinates was provided by MSU Geography Department (Courtesy of Dr Andresen and Dr Gopal). This Climate data was generated using the funding support from the project Dynamic Interactions among People, Livestock, and Savanna Ecosystems under Climate Change, Award No. BCS/CNH 0709671, from the National Science Foundation (NSF) Bio-complexity of Coupled Human and Natural Systems Program.

amounts below 40mm to capture moisture stress. This variable is used together with rainfall amounts in the estimation process.

The production function also includes a dummy for single-cropping to delineate mixed crop plots from single crop plots. We assume some systematic differences in yields between these plots and will test for it. A dummy for tractor or ox-plough land preparation is also included against manual hoe.

BIBLIOGRAPHY

BIBLIOGRAPHY 1

- Abrar, S. and Morrissey, O. (2006). Supply Response in Ethiopia: Accounting for Technical Inefficiency. *Agricultural Economics*, International Association of Agricultural Economists, vol. 35(3), pages 303-317, November
- Alene, A. D., Manyong, V.M., Omany, G., Mignouna, H.D., Bokanga, M., and Odhiambo G. (2008). Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *Food Policy*, Volume 33, Issue 4, August 2008, Pages 318-328
- Alvarez, A., Amsler, C., Orea, L., and Schmidt, P. (2006). Interpreting and Testing the Scaling Property in Models where Inefficiency Depends on Firm Characteristics. *Journal of Productivity Analysis*, Springer Netherlands. 25: 201-212.
- Allgood, J. H., and Kilungo, J. 1996. *An appraisal of the fertilizer market in Kenya and recommendations for improving fertilizer use practices by smallholder farmers: A field report*. Muscle Shoals, Ala., U.S.A.: International Fertilizer Development Center (IFDC).
- Ariga, J.M., And Jayne, T.S. (2010). Factors Driving the Increase in Fertilizer Use by Smallholder Farmers in Kenya, 1990-2007 in *Yes Africa Can: Success Stories From A Dynamic Continent*, World Bank Report.
- Baanante, C.A. (1997). "Economic Evaluation of the Use of Phosphate Fertilizers as a Capital investment" paper in A.E Johnston and S.K Syers (eds): *Nutrient Management for Sustainable Food Production in Asia*. CAB International Publishers, Wallingford, UK.
- Chamberlain, G. (1984). Panel data, Chapter 22 in Z. Giliches and M. Intrilligator,, eds., *Handbook of Econometrics*, North-Holland, Amsterdam: 1247 - 1318.
- Chhibber, A. (1989). The Aggregate Supply Response: a Survey. in S. Commander (Ed), *Structural Adjustment and Agriculture*, London: ODI.
- Crawford, E., Kelly, V., Jayne, T. S., and Howard, J.(2003). Input use and market development in Sub-Saharan Africa: an overview. *Food Policy Special Issue on Input Use and Market Development in Sub-Saharan Africa* 28(4): 277-292.
- IFDC. 2001. *An assessment of fertilizer prices in Kenya and Uganda: Domestic prices vis-à-vis international market prices*. International Center for Soil Fertility and Agricultural Development (IFDC), Muscle Shoals, Alabama, U.S.A. Photocopy.
- Jayne, T. S., and Argwings-Kodhek, G. (1997). Consumer response to maize market liberalization. *Food Policy* 22 (5): 447-458.

- Kenya Broadcasting Corporation. (2010). Farmers to receive subsidized fertilizers.
<http://www.kbc.co.ke/story.asp?ID=62799>. March 17
- Kimuyu, P. 1994. *Evaluation of the USAID/Kenya fertilizer pricing and marketing reform program*. Nairobi, Kenya: U.S. Agency for International Development / Kenya, Nairobi.
- Liu, Y. and Myers, R. (2009). Model selection in stochastic frontier analysis with an application to maize production in Kenya. *Journal of Productivity Analysis*, Springer Netherlands. 31: 33-46.
- Marenja, P. P. and Barrett, C. B. (2009). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics* 40(5): 561-572.
- Ministry of Agriculture. (2008). *Economic review of agriculture: 2008*. Central Planning and Project Monitoring Unit. Nairobi, Kenya: Ministry of Agriculture.
- Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica* 46(1): 69-85.
- Ogbu, O.M. and Gbetiyou, M. (1990). Agricultural Supply Response in Sub-Saharan Africa: A Critical Review of the Literature. *African Development Review* 2: 83-89.
- Phimister, E. and Roberts, D. (2006). The Effect of Off-farm Work on the Intensity of Agricultural Production. *Environmental and Resource Economics*, Springer Netherlands. 34: 493-515.
- Savadogo, K., Reardon, T., Pietola, K. (1994). Farm productivity in Burkina Faso: effects of animal traction and non-farm income. *American Journal of Agricultural Economics* 76: 608–662.
- Staatz, J. M., Dione, J., and Dembele, N. N. (1989). Cereals market liberalization in Mali. *World Development*, Elsevier, vol. 17(5), p:703-718
- Suri, T. (2005). Selection and Comparative Advantage in Technology Adoption. Department of Economics, Yale University / Tegemeo Institute.
- The Daily Nation Newspaper at <http://www.nation.co.ke/News/-/1056/836926/-/vorgt3/-/index.html>. 01/06/2010.
- Wanzala, M., Jayne, T. S., and Staatz, J. 2002. *Fertilizer markets and agricultural production incentives: Insights from Kenya*. Working Paper 4. Tegemeo Institute. Nairobi, Kenya: Egerton University.
- Weber, M. T., Staatz, J.M., Holtzman, J.S., Crawford, E.W., and Bernstein, R.H. (1988).

Informing Food Security Decisions in Africa: Empirical Analysis and Policy Dialogue. Invited Paper Session on Economic Policy Adjustments in Developing Countries, AAEA Annual Meetings, Knoxville, Tennessee, July 31 to August 3, 1988. Michigan State University Department of Agricultural Economics Staff Paper No. 88-58.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel data*. Cambridge, MA: MIT Press, 2002.

World Bank Country Reports for various years. Kenya.

Yanggen, D., Kelly, V., Reardon, T., and Naseem, A. (1998). Incentives for Fertilizer Use in Sub-Saharan Africa: A Review of Empirical Evidence on Fertilizer Response and Profitability. Michigan State University, Department of Agricultural Economics, International Working Paper No. 70.

ESSAY 2:

ESTIMATING DEMAND FOR FERTILIZER: INCORPORATING ENDOGENEITY OF PRIVATE INVESTMENT IN RETAIL MARKETS

2.1.2 Introduction: Importance of Estimating Fertilizer Demand and Profitability

The sharp rise in fertilizer prices starting in 2007 has been attributed to increased demand for fertilizer resulting from expansion in bio-fuels production, use of increasingly expensive petroleum and natural gas inputs in the manufacture of fertilizers, and implementation of protectionist policies by countries like Russia and China aimed at reducing fertilizer exports in a bid to aid domestic production (Zhao, 2009). By 2009 average prices for Di-Ammonium Phosphate fertilizers had risen by more than 100 percent compared to 2007 levels. This has policy implications for Sub-Saharan Africa (SSA) which has the lowest fertilizer application rates in the world at 9 kilograms per hectare (IFDC 2006), a mere ten percent of the world's average and just five percent of East Asia's average.

Further, the current global economic downturn may reduce effective demand for farm inputs which has implications for food security in developing countries. In order to raise the low productivity for smallholder farmers, one of the major thrusts of developing economies is to encourage increased input use and intensification through subsidies and input market liberalization efforts to raise farm productivity¹¹. Such activities include provision of free fertilizer and seeds to poor households. The return of state market-participation in form of subsidies can be explained by some frustration with stagnant agricultural growth even after the

¹¹ For instance the Kenya government's Agricultural Sector Program Support (ASPS) targets investment in inputs in order to increase productivity and rural incomes.

liberalization of markets, increased poverty levels, and the need to boost effective demand for producers to improve national food security.

However, a number of studies on the performance of subsidy programs provide no discernible consensus (Minot 2004; Morris et al 2007; Crawford et al 2006). While some studies call for an African “green revolution” by increasing the role of the state in markets (Dorward et al 2004) or fostering policies that lessen economic and market constraints that impede fertilizer use (Gregory, 2006; Jayne et al., 2003; Omamo et al., 2002; Poulton et al., 2006) others suggest that this effort to correct market failures may “crowd out” the private sector participants which then provides credence for more state involvement (Jayne et al 2003) therefore replacing commercial systems with government-run programs, a situation that evokes the pre-liberalization period.

Some studies argue that soil quality has deteriorated in SSA and therefore fertilizer intensity is part of the solution. Ange(1993) predicts declining fallow land, and FAO(2000) reveals increasing deforestation, Scherr (1999) show increasing degradation due to human activity, and others indicate negative N, P, K balances(Pol 1992). However, these studies do not analyze the effect of market forces like fertilizer and output prices and how these create incentives. Nor do these studies utilize econometric methods to model relationships between variables of interest.

Some articles study the effect of household characteristics, infrastructure, and transaction costs on fertilizer demand (Jayne et al., 2003; Omamo et al., 2001; Diagne and Zeller, 2001). Others look at role of government policy (exchange rate, prices, and subsidies), human capital development, credit, market information, and regulatory mechanisms in influencing demand for fertilizer (Kelly, 2006; Crawford et al, 2006). Though this essay, like a number of other studies,

includes distance to markets as an explanatory variable, it is however treated as potentially endogenous and an instrumental variable approach used to avoid possible bias.

The transaction cost studies argue the case of missing markets that result in households preferring autarky or subsistence in the presence of uncertain markets. Adesina (2002) finds that distance to the market has a negative effect on adoption of fertilizer. However, the author does not check whether this variable is exogenous or not which could bias results. Barrett (2009) argues that in addition to transaction costs, households with productive assets have a higher probability of participating in agricultural markets, with these assets (land, animals, etc) providing a form of insurance and liquidity against shocks. Others look at the role of micro-finance institutions and credit in reducing the problem of low fertilizer demand by tackling liquidity constraints (Kherallah et al., 2002; Omamo et al., 2001). Duflo et al(2007) use experimental data from western Kenya to conclude that providing information to producers on optimal rates and making sure that fertilizer is available in stores right after harvest when crop sales might provide the necessary cash, will raise uptake. Croppenstedt and Muller (2000) estimate fertilizer demand in Ethiopia using a double hurdle model and find that the level of education for the head of household is significant for fertilizer intensity and credit, household size, and the value-to-cost ratio have a significant effect on both adoption and intensity of fertilizer use.

In Essay 1 we produce marginal value-to-cost ratios from a primal production function, generate the average contribution of fertilizer to household income, and then compare trends in this latter variable across different locations and households. The heterogeneity in the contribution from nitrogen fertilizer across location and districts as shown in Essay 1 can be explained by results of Essay 2.

2.1.3 Objectives

This essay provides a number of key contributions to policy debate and the literature on fertilizer productivity and food security. While Essay 1 generates estimates of fertilizer contribution to income in different parts of the country using a production function, Essay 2 provides insights into what influences the demand for fertilizer (both the purchase decision and rate of application per acre) and helps to explain the heterogeneity in fertilizer use across the country. Jointly, these two essays then will shed light on why demand may be less than optimal even when value-to-cost ratios are attractive enough to trigger more intensification of fertilizer use. This will lead to identifications of constraints to increased fertilizer use.

A key objective in Essay 2 is to measure the contribution of reduction in distance from farm to the nearest fertilizer seller to demand for fertilizer. We use distance to nearest fertilizer seller as a proxy for private trade investments in retailing and estimate its effect on fertilizer demand. Figure 11 and Table 11 shows that this variable has been declining since the markets were liberalized. Increased use of fertilizer has gone hand in hand with increased retailing activities and so some of the factors that influence demand for fertilizer and which we are unable to capture in our model are likely to influence retail activities as well. The estimation of demand without accounting for possible lack of exogeneity in this distance variable may bias our results. Factors that influence demand for fertilizer but that are not controlled for in our model in Essay 2 (either because they are missing or not collected by survey) may probably also be correlated with distance to fertilizer seller. In this study we test for exogeneity and employ instrumental variable approach to mitigate bias in the estimates for fertilizer demand.

In order to get reliable estimates we need to apply appropriate econometric methods that control for possible endogeneity and unobserved heterogeneity across households. This essay

uses distance from farm to nearest fertilizer seller as a proxy for expansion in fertilizer retail services. First, we test for exogeneity of distance to fertilizer seller and then use instrumental variables to estimate demand for fertilizer. This is an important step in mitigating bias which similar studies that have used distance as an explanatory variable have neglected. It is possible that some of the factors affecting fertilizer demand whose data was not captured in the panel (and so are included in the error term) also influence the dynamics in distance from the farm to nearest fertilizer seller. For instance Constituency Development Funds (CDFs) that were distributed by the central government to constituencies for infrastructure projects could influence demand as well as distances to sellers¹². Use of cross-sectional data suffers from biased and inconsistent estimates but with panel data we model unobserved heterogeneity by tracking each household across the period using relevant econometric techniques. This essay uses panel data Correlated Random Effects (Mundlak, 1978; Wooldridge 2002) that assumes arbitrary correlation between unobserved heterogeneity and explanatory variables in order to avoid bias.

This essay makes methodological as well as informational contributions to existing literature. By documenting the effect of various factors that have contributed to declining distance between farms and fertilizer retailers, the study provides important information to governments on private sector incentives aimed at attracting businesses to rural areas. With governments in developing countries encouraging private sector fertilizer trade, this information will provide insight on what should be done to boost fertilizer sales in rural areas. Input retailers are assumed to invest in areas where demand is likely to be relatively higher. Results from this study also contribute towards targeted policy aimed at specific regions or type of farmer based

¹² These funds were used for various projects at different locations- for construction of roads, school, water projects and clinics. Accessibility, created by roads could encourage increased fertilizer use by farmers while at the same time sellers could invest closer to farms if roads were in good condition.

on unique characteristics. Some agricultural policies in developing countries are too broad in nature, prescribing the same options for producers or consumers who are vastly heterogeneous. For instance, fertilizer application rates for Kenya maize are recommended at 100 kilograms of basal and top-dressing per acre regardless of local agro-conditions. These recommendations are based on studies done in the 1970s on soil quality but these conditions have changed since due to nutrient-mining and erosion among other factors. We control for and estimate agro-climatic, regional, and household differences in demand for fertilizer.

2.1.4 Modeling Framework

This essay models a number of economic relationships. The first relationship looks at factors that explain why producers buy fertilizer or not. The second looks at factors that influence the quantity of fertilizer that farmers apply per acre. Economic theory posits that prices are the key variables that drive behavior and this essay controls for this but also includes non-price factors that also influence producer decisions. These latter include household demographics, socio-economic factors, agro-ecological conditions, and market access factors like distance to the nearest fertilizer seller.

If we simply express $x^*(p, w; Z)$ as the fertilizer demand relationship then p, w represent prices of output and inputs while Z stands for the other factors mentioned above. The Z includes distance to the nearest fertilizer seller which we take to be endogenous and do relevant tests and search for appropriate instrumental variables (IVs). As explained in the data and variable sections below, distance to seller is used in this study as a proxy to measure the expansion of fertilizer retail services. The distance from farm to fertilizer seller has been declining over the panel period implying geographic growth and diversification in retailing. As

part of the process of testing for endogeneity and looking for IVs, we estimate an auxiliary relationship that provides important information on what is driving retailers to move closer to farms.

Under the above framework, we tackle the issue of unobserved heterogeneity by incorporating the Mundlak-Chamberlain approach in a Correlated Random Effects (CRE) context that assumes arbitrary correlation between the unobserved heterogeneity and explanatory variables.

2.1.4.1 Double Hurdle and Correlated Random Effects

We treat the household's decision to participate in fertilizer markets as a two-part procedure involving the decision on whether to participate in the fertilizer market as a buyer and how much quantity to use per acre (Goetz, 1992; Key et. al., 2000; Bellmare and Barrett, 2006; Cragg, 1971). The two decisions need not be independent and tests can be conducted to determine if the second decision is conditional on the first. If this is the case then both equations need to be estimated jointly. We use the double hurdle conceptual framework to estimate the demand for fertilizer. The double hurdle model is commonly used with data consisting of a significant number of zeros for the dependent variable (Haines et al 1988). The model consists of decision to purchase fertilizer and quantity consumed equations, where P_{it} is the probability of purchasing fertilizer by household i at time t and Y_{it} is the amount of fertilizer applied per acre:

$$P_{it} = \begin{cases} 1 & \text{if } p_{it}^* > 0 \\ 0 & \text{Otherwise} \end{cases} \quad \text{and} \quad Y_{it} = \begin{cases} y_{it}^* & \text{if } y_{it}^* > 0, p_{it} = 1 \\ 0 & \text{Otherwise} \end{cases}$$

Where the probability of purchasing equation can also be expressed by the indicator function

$P_{it} = 1[p^*_{it} > 0]$; $p^*_{it} = Z_{jit} \alpha + V_{it}$ is a latent, unobservable variable that explains purchase decision; and $y^*_{it} = x_{jit} \beta + u_{it}$ is the latent variable for Y_{it} . The subscripts i denote household, j variable, and t panel year respectively.

Y_{it} = the observed fertilizer application rates per acre for household i at time t

Z_{jit} , x_{jit} = vectors of explanatory variables accounting for purchase and use, respectively. These vectors are assumed to be exogenous (except for distance to seller) and they need not have the same variables.

V_{it}, u_{it} = error terms that are independent and distributed as $V_{it} \sim N(0, \sigma_v)$

and $u_{it} \sim N(0, \sigma_u)$.

$\alpha, \beta, \sigma_v, \sigma_u$ = parameters to be estimated

The distributional forms of the separate stochastic processes that determine purchase and consumption can be expressed in terms of the normal distribution:

$$F(P_{it}=1 | Z_{jit}) = \Phi((Z_{jit} \alpha) / \sigma_v) \text{ and}$$

$$F(P_{it}=0 | Z_{jit}) = 1 - \Phi((Z_{jit} \alpha) / \sigma_v)$$

$$f(y_{it} | x_{jit}, P_{it}=1) = \frac{\phi((y_{it} - x_{jit} \beta) / \sigma_u)}{\sigma_u \Phi((x_{jit} \beta) / \sigma_u)}, y_{it} \geq 0$$

F , f , Φ , ϕ stand for probability, conditional density function, cumulative distribution function, and probability density function respectively. Therefore, assuming normal errors, the first equation corresponds to the “probit” model. The second hurdle takes the form of a “tobit” model that can handle zero levels of fertilizer application rates independently of the first hurdle. It is further assumed that the latent variables have a bivariate normal distribution:

$$v_{it}, u_{it} \sim \text{BVN}(0, \Sigma), \Sigma = \begin{bmatrix} \sigma_u^2 & \rho \\ \rho & \sigma_v^2 \end{bmatrix}$$

As Blaylock and Blissard (1992) point out, this general model nests a number of other formulations. For instance when $\rho = 0$ the model collapses to the independent Cragg model where the processes are independent. The Tobit model is nested within the independent double hurdle model ($\rho = 0$) when it is further assumed that the probability of purchasing is one. The double hurdle model has two advantages over the tobit model. The latter interprets the zeros on amount of fertilizer as corner solutions i.e. the household chooses not to purchase any amount at the prevailing levels of exogenous variables, such as prices i.e. both decisions are determined by a single mechanism. A further restriction of the Tobit is that both the decision to purchase fertilizer and the amount of fertilizer applied per acre are determined by the same variables, and the assumption that a variable has the same magnitude and direction in both purchase and level equations (Wooldridge 2002). For instance, this latter assumption is not necessarily true for level of education in labor participation and hours worked model where education level may have different magnitude and sign in the two models.

2.1.4.2 Modeling Unobserved Heterogeneity

The problem of unobserved heterogeneity (resulting from measurement errors or missing variables whose data has not been collected like managerial ability, soil quality, and other unobserved time-constant factors) needs to be accounted for to avoid inconsistent estimates. The second equation that estimates the quantity of fertilizer applied per acre can be expressed as follows (the purchase equation can also be written similarly) when modeling for unobserved heterogeneity:

$$y_{it} = \beta_0 + \sum_{j=1}^k \beta_j x_{jit} + c_i + e_{it} \quad (1), \text{ where } u_{it} = c_i + e_{it} \quad (\text{error } u_{it})$$

includes unobserved heterogeneity, c_i). The issue of unobserved heterogeneity is evident if $\text{cov}(x, c_i) \neq 0$ and c_i is part of the error structure, which makes pooled ordinary least squares (OLS) biased and inconsistent for instance. If c_i is present and influences each year's optimal input and output decisions it needs to be controlled for else feedback might exist (Hausman, Hall, and Griliches 1984) which means no strict exogeneity.

For this essay we model unobserved heterogeneity using the correlated random effects model (CRE) under strict exogeneity but assume arbitrary correlation between c_i and time-varying exogenous factors¹³ (Mundlak 1978; Chamberlain, 1984; Wooldridge 2002). This approach replaces c_i with its projection onto the explanatory variables including the projection

¹³ Using fixed effects (FE) produces incidental parameters problem (Wooldridge 2002; Heckman 1981) when dealing with fixed and small T with large N which leads to inconsistent estimates for β when estimating c_i as it becomes difficult to integrate out c_i in the likelihood function.

error term and assumes no particular distribution for $E(C_i | X_i)$. The CRE approach, unlike RE and FE, models a relationship between C_i and X_{jit} or Z_{jit} , using a Mundlak-Chamberlain approach;

$C_i = \tau + \bar{X}_i \xi + a_i$ where $C_i \sim \text{Normal}(\tau + \bar{X}_i \xi, \sigma_a^2)$, where $\bar{X}_i$ is the average of X_{it} , $t=1, \dots, T$, for each unit i.e. $\bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}$ for $i=1..N$. Also τ and ξ are parameters, and σ_a^2 is the time-invariant variance for a_i assuming $a_i | X_i \sim \text{Normal}(0, \sigma_a^2)$ and $\text{cov}(u_{it}, a_i) = 0$ (Wooldridge 2002). This equation is plugged into the main regression equation (1) to replace C_i . In the actual estimation process τ is absorbed into the overall intercept or constant. Using a Wald test the hypothesis of no correlation between C_i and X_i i.e. $H_0: \xi = 0$ can be tested in order to make a choice between CRE and other approaches (pooled model and random effects model that assumes independence between C_i and X_i).

The CRE approach is useful in a number of ways; first if the null $H_0: \xi = 0$ is rejected, this implies that the explanatory variables are correlated with unobserved heterogeneity. Then the estimation process favors FE over RE which assumes that C_i are time-constant and so can be eliminated by the within estimation differencing process. This test is similar to the Hausman

test between RE and FE. Secondly, the CRE approach provides an intuitive way of estimating changes that occur “within” the panel unit (households, individuals, firms) over the period and measures of differences “across” units. Since $\bar{X}_i$ is constant over time, the ξ parameters in the structural equation shed light on what differences exist across households or units while β provides a measure of changes “within” the households. An alternative way to lessen the effect of unobserved heterogeneity is to include dummies for groups or regions (villages, etc) that may be correlated with time-constant variables in X_{it} , the explanatory variables, and then use Random Effects to account for any residual serial correlation resulting from unobserved time-constant variables.

2.1.4.3 Testing for Endogeneity of Distance to Fertilizer Retail Services

One of the hypotheses to be tested is whether distance from farm to nearest fertilizer seller is exogenous to fertilizer demand. If this is endogenous then its correlation with factors that influence fertilizer demand but are not included in the model will lead to inconsistent estimates, since $E(X'U)=0$ fails where U is the error component in equation (1). One option to deal with endogeneity is to use lagged distance to fertilizer seller assuming that it is uncorrelated with the current unobservable variables. This leads to loss of one period’s data and lagging might also pose a problem with assumption of strict exogeneity where the error term needs to be uncorrelated with explanatory variables in all periods (Wooldridge 2002) i.e. no feedback. Other possibilities include identifying location-specific and household factors that are correlated with distance but that do not relate to error term components.

Sing M-C device with such variables may help mitigate endogeneity. A potential instrumental variable (IV), say V , has to meet two requirements;

a) $\text{Cov}(V' U)=0$ i.e. exogenous in the demand equation, and

b) $\text{Cov}(V' X_e) \neq 0$ where X_e is the endogenous variable.

Condition a) cannot be tested since population U is unknown. In order to figure out b) we project the endogenous variable onto all exogenous variables including the potential IV variable in a reduced form equation or auxiliary or control equation to estimate the partial correlation between V and X_e after netting out all other variables (Wooldridge 2002). For simplicity the following structural and reduced form equation have been stripped of individual and time sub-scripts.

$y = \alpha + \beta_1 x_1 + \dots + \beta_k x_k + \beta_e x_e + U$, with X_e being endogenous in the sense

$E(X_e U) \neq 0$ and $X_e = \delta + \beta_1 x_1 + \dots + \beta_k x_k + \beta_{iv} V + u_X$ is the reduced or control equation

The variables y , X_e , and V are fertilizer quantity, distance to fertilizer, and instrument variable respectively. After estimating the reduced (control) equation the predicted residuals $\hat{u}_X$ are plugged back into the structural equation

$y = \alpha + \beta_1 x_1 + \dots + \beta_k x_k + \beta_e x_e + \phi \hat{u}_X + \mathcal{E}$, $\mathcal{E}$ is the random error term.

If the null of $H_0: \phi = 0$ is rejected, then X_e is not exogenous and we need to use appropriate IVs to estimate the demand for fertilizer. In case there is more than one IV for each endogenous variable, it is important to test for over-identifying restrictions.

2.2 Empirical Model

Combining the approaches described in the modeling framework above, we have three equations to estimate; the demand for fertilizer in a double hurdle framework that incorporates the Correlated Random Effects (CRE) approach for unobserved heterogeneity and possible endogeneity of fertilizer retail services (distance to nearest fertilizer seller).

$$p_{it} = \delta + g_{ijt} \theta + d_{it} \tau + \bar{x}_i \lambda + r_{it} \quad (\text{decision to purchase fertilizer})$$

$$y_{it} = \alpha + x_{ijt} \beta + d_{it} \gamma + k_{it} \varsigma + \bar{x}_i \xi + \mu_{it} \quad (\text{fertilizer kilos per acre})$$

$$d_{it} = \eta + h_{ijt} \Re + v_{it} \omega + e_{it} \quad (\text{distance to nearest fertilizer seller}),$$

where r_{it} , μ_{it} , and e_{it} are error terms with the usual assumptions of zero expectation.

p_{it} = fertilizer purchase by household i at time t (0,1)

y_{it} = quantity of fertilizer in kgs applied per acre by household i at year t

x_{ijt} = indicates value of factor j for household i at time t ; these include fertilizer to

maize price ratio, distance to road, distance to extension service, asset values, age,

education level (yrs), land preparation technology (manual, oxen, tractor), rainfall

amounts, indicator of moisture stress, soils types. Note that g_{ijt} and

h_{ijt} may have common variables.

d_{it} =distance from farm household i to fertilizer seller at time t

$\overline{x_i}$ =average of time-varying x_{ijt} , d_{it} , and k_{it} for household i

k_{it} =dummy variables for gender, agro-ecological zones, land preparation

technology for household i at time t .

v_{it} =Instrumental variable for d_{it}

δ , α , η =intercepts

θ , β , $\Re$, τ , γ , ω , ς , λ , ξ = vector of parameters to be estimated

The third equation is a reduced form (control or auxiliary) equation of the endogenous variable on all exogenous variables, including possible IVs. If the null $H_0: \omega = 0$ is rejected at 5% or 1% significance then V is a valid IV for d (Wooldridge 2002). The IV variable V has to be from outside the pool of all exogenous variables in the model i.e. not part of the existing exogenous variables included in the estimation (the exclusion restrictions).

It is easy to test whether the variable is exogenous by inserting the estimate $\hat{e}_{it}$ from the reduced form equation into the other equations and testing whether the coefficient is

significant (Wooldridge 2002), which incidentally provides 2SLS estimates as well. For a single IV case it can be shown that τ and γ , the parameters for distance in the decision to purchase fertilizer and consumption equations above are identified i.e. there is enough information to estimate these parameters from the sample.

2.3 Data Description and Methodology

Data used is described more comprehensively in sub-section 1.1.3.1 and A1.2 in Appendix 1. The description of variables used in Essay 2 is given in Section 2.3 and Table 11 to Table 17 in Appendix 2. We use data from Tegemeo Rural Household Survey for the years 1997, 2000, 2004, 2007. The rural survey, funded by USAID, has been a joint effort between Tegemeo Institute (Egerton University, Kenya) and Michigan State University. We focus on plots that are planted with maize because maize is a major food crop grown by over 95% the sample households and accounting for over 50% of fertilizer use (Ministry of Agriculture Annual Reports, Tegemeo Household Survey). The data covers mostly production data for major crops (collecting inputs, outputs, and price data) and off-farm activities.

There two planting seasons, major and minor, with major season accounting for over 80% of annual maize output. Most households have two maize crops in a year, one during the longer rain period and the other during the shorter rain season. Short season rainfall maize seed is composed of quick-maturing varieties compared to long season.

In addition to the seasonal nature of production there are two distinct geographic regions with vastly different agro-ecological conditions. The low potential region covers the lowland zones in the eastern and coastal areas while the high potential region consists of the mid and highlands areas in central, Rift valley, and western parts of the country. The latter region has more rainfall,

better soils for maize production, more infrastructures (roads, electricity, and schools) probably due to its potential and productivity which attracts more investments. Some of these factors might also explain the concentration of fertilizer retailers across the country. There is also more complete data available in this region compared to the low potential region.

2.3.1 Definition and Summary Statistics of Variables

The definition of the variables used in this essay and their summary statistics at different percentiles are given in Table 11 to Table 17 in Appendix 2. This essay consists of two regressions that capture fertilizer demand conditions for Kenya maize. The first regression provides estimates of factors that affect farmers' choice to buy (or not buy) fertilizer while the other equation looks at factors that influence their decision on the amount of fertilizer (in kilograms) to apply per acre. Although there is heterogeneity across the region, average mean fertilizer application rates rose from 44 kilos in 1997 to 57 kilos per acre in 2007, a twenty seven percent increase. For explanatory variables we include the following: nitrogen fertilizer-maize price, values of household assets (in Kenya Shillings), distance to nearest fertilizer seller in kilometers, distance to motorable road (kilometers), distance to extension service in kilometers, education of head of household (number of years of school), gender of household head, age of head of household (years), total land owned (acres), land tenure dummy (owned with title, owned without title documents, rented), dummies for agro-ecological conditions, rainfall (in mm), moisture stress during the growing season(measured as the proportion of twenty-day periods that had mean rainfall below 40mm per day), and soil types. Rainfall data was simulated at the village level using a combination of data from the nearest weather stations and satellite imagery.

The distance from farm to extension services, to motorable road, and to fertilizer seller were collected at the household level. Distance to fertilizer seller has been declining over the period of the panel survey. The same factors that influence demand for fertilizer (relative prices, wealth, weather, soils, roads, education levels) may also influence traders in their decision on which areas to target with their products. Table 11 shows a declining trend in distance between the farm and fertilizer sellers for the panel period.

One of the expected outputs is to estimate the effect of factors that have contributed to this trend which will provide useful information for policy aimed at encouraging increased private sector investment in fertilizer distribution. Identifying factors that have created incentives for traders to move closer to farms will enable policy makers to put in place plans that can augment private sector efforts.

The formal education level of the household head is expected to have a positive effect on demand both probability and amount of fertilizer used. Data for this variable was not collected for the year 1997. Education level does not change significantly across the panel period and this may be attributable to most heads of households not advancing their formal education during the duration of these surveys. So the variable will provide comparisons across households rather than within household dynamics. Female headed households are expected to consume less fertilizer than their male counterparts. We assume that gender bias exists in resource allocation in these communities.

The model includes the value of agricultural assets (value of draught animals and equipment like ploughs and sprayers) and total acres owned by household to represent wealth. We expect a positive relationship between fertilizer demand and wealth as access to funds to purchase more fertilizer becomes available. We also use dummies for types of land ownership

(land tenure categories include those who own land and have title to that land, own but no title, and renters). The nitrogen-maize output price ratio is expected to be negative, implying that relatively higher increase in nitrogen prices compared to maize prices will lead to decreased demand for fertilizer. Soil and climatic factors are expected to influence demand for fertilizer; we include soil types, rainfall, and agro-ecological zone dummies. Note that these last three factors may be very collinear and particularly the soil type and ecological factors (rainfall varies across villages while soil and zones vary across districts and zones).

2.4 Results

2.4.1 Factors Influencing Distance to Fertilizer Seller

As a precursor to the estimation of fertilizer demand below we estimate the effect of various factors on the distance to fertilizer retail seller in order to identify possible instruments and also to provide predicted residuals that will be used to test for endogeneity in the structural equation. It is important to understand what “distance from farm to fertilizer seller” really stands for. This is an estimate from producers on what they perceive is the distance in kilometers from their farm to the location of the nearest fertilizer seller. Typically most smallholders buy their fertilizer from medium stores in the nearest town center and small grocery stores in the neighborhood that also sell fertilizer during the farming season depending on the quantity of fertilizer they need. So for such farmers distance is basically the number of kilometers from their farm to where they buy fertilizer.

Therefore in this section we try to understand what factors drive the proximity to retail seller. The regression of distance to fertilizer seller (kilometers) reveals some interesting results. We run pooled OLS, fixed effects (FE), and correlated random effects (RE) as shown in Table 20. The econometric assumptions for these approaches and model specifications are dealt with in more detail under discussion of the modeling framework in Essay 1.

In the following sections we discuss the various results shown in Table 20. Results for all these models have similar signs and significance levels for most of the estimates. The estimates for the price ratio (nitrogen-maize) indicate that households facing higher fertilizer prices also deal with longer distances to fertilizer delivery services. However, for the low potential region the signs are reversed probably because there are very few fertilizer users and the prices that are available cover large areas i.e. limited variation. The effect of distance to extension service and

distance to motorable road are positive and significant. This is not surprising as extension service provision and transport costs are factors that fertilizer dealers take into consideration in areas where they plan to invest. Probably fertilizer retailers' role as input providers is complemented by presence of extension service being nearby.

Households with higher wealth measures (land size and value of agricultural assets) are relatively further away from fertilizer services. This may be a consequence of the trend in most rural town centers where business activities (grocery stores, milk vendors, shoe shiners and repairers, and other small businesses) tend to locate near high population centers which happens to consist of small farms clustered together. These are also areas where government extension workers have their offices in order to reach as many smallholders as possible. In general, households with older heads are further away from fertilizer sellers than those with younger heads. The data shows that households that are more than six miles from a fertilizer seller have an average education level one year less than their counterparts but they own nearly twice as much land and assets.

The amount of rainfall has a negative and significant effect on the distance to fertilizer seller relative to the farms. Villages with higher rainfall amounts are located closer to fertilizer delivery points. On the other hand areas with relatively higher moisture stress during the maize growing season are relatively further away from retail services. These variables imply that rainfall (water availability) is important in the getting access to fertilizer sellers. The agro-ecological zones and soils can also be explained similarly. Soils are described in Table 10 and zones in Section A1.2. Households in zone 1 are relatively further away from fertilizer sellers than those in zone 2 and 3 for low potential region. Similarly zone 4 has relatively longer distances to sellers compared to zones 5-7 for the high potential region.

Therefore, household wealth, soil conditions, rainfall, education level of the head of the household, and moisture stress influence the relative distance from farms to the nearest fertilizer seller i.e. retail fertilizer growth. We identify number of years the household is resident in the area as a potential instrumental variable for distance to fertilizer seller in estimation of fertilizer demand below. This is highly correlated with distance to seller (explained more below).

2.4.2 Estimating Demand for Fertilizer in the Context of Endogeneity

In this section we estimate demand under two assumptions, existence of unobserved heterogeneity and the assumption that the distance to fertilizer seller may be endogenous in our structural fertilizer demand model. First we use a two-stage process as explained in Wooldridge (2002) to test for endogeneity of retail services. This is simply plugging the estimated residual from the auxiliary regression for the endogenous variable in our structural model and testing for significance. A rule suggested by Wooldridge is that rejecting the null of exogeneity at below 5% level is strong enough to proceed with instrumentation in order to avoid inconsistent estimates. We reject the null at close to 5% level which indicates that distance to retail services may not be exogenous in our fertilizer demand model. However, because testing for exogeneity is not always straight forward and depends on choice of model and variables that one uses as instruments, we proceed in a double-pronged way by using methods that assume exogeneity and comparing with results from methods that assume endogeneity.

We use a double hurdle model approach to estimate the equations on the decision to buy and the amount of fertilizer to apply per acre using Burke (2009) STATA commands. Both the decision to buy and amount of fertilizer to apply will be estimated using panel random effects GLS estimators. In both equations we incorporate the Mundlak-Chamberlain device for

unobserved heterogeneity. Using CRE approach mitigates the possible endogeneity of retail services by controlling for unobserved effects that contribute to the endogeneity.

Another possible method is the Tobit which uses censoring techniques and assumes both regressions are determined by the same process and the estimates have same magnitude and signs for both equations. This model is nested in the double hurdle model. Though we run the “tobit” model, it is not reported since it is too restrictive in its assumption about the data generating process being the same for both purchase and intensification decisions.

We reject the null of no unobserved heterogeneity which implies that fixed effects approach is the appropriate method over random effects that assumes exogeneity. CRE offers the benefits of producing the same estimates as those from a FE regression while at the same time providing a way to model heterogeneity so as to get estimates for time-invariant variables that might otherwise be differenced away by the FE approach.

Therefore we discuss results for the CRE method and compare with pooled OLS and random effects instrumental variable approach. The random effects instrumental variable approach uses two-stage least squares to estimate models in which some covariates are endogenous. Therefore we compare instrumental approach (controls for exogeneity) with CRE (where we assume exogeneity). All these methods show similar signs for most of the estimates, some indication of robustness.

Though the estimates from IV approach have same sign and significance levels as those from CRE model, they are slightly larger in magnitude (Table 23 and Table 24). However, a Hausman test for systematic differences between the IV (assume consistent) and CRE (assume efficient under the null of exogeneity) estimates gives a p-value of 0.708, which means we reject the null of H_0 : (difference in coefficients not systematic) at 8%. Endogeneity has the potential of

resulting in biased estimates depending on the nature of the relationship between the explanatory variables and missing variables that are lumped with the error term.

From our regressions on distance to fertilizer seller we identify the length of time that the household has lived in the current location as an appropriate IV. It is correlated with distance to seller (Table 16 and Table 17) though we cannot confirm the second requirement of exogeneity with error term since this is a population measure (we can only use sample analogues whenever possible but we have no information as to the distribution of the error structure). We can only make tangential arguments for the case that longevity of the household in the current location is not correlated with missing factors that influence demand for fertilizer.

We discuss both results from the first and second regression equations on probability of purchasing fertilizer and quantity of fertilizer applied per acre (Table 21 to Table 24) because these results are cross-cutting in their ramifications. Table 21 and 22 contain regressions on probability of purchasing fertilizer using both panel random effects probit and probit on pooled data. For both the panel and pooled probit results, the significance levels and signs of estimate are similar.

The distance to a fertilizer seller has a statistically significant negative effect on probability of buying fertilizer (Table 21) but has no significant influence on the quantity in kilograms of fertilizer applied per acre (Table 23) for the high potential region. However, for the low potential region the sign is negative and significant for both decisions. A household that is located 10 kilometers away from fertilizer retailers has 0.1 lower probability points of purchasing fertilizer for the high potential region but 0.2 points lower for the low potential region. This implies that even though proximity to retail services does influence the decision to purchase fertilizer, the amount of fertilizer applied per acre does not change significantly with distance for

the high potential region but it does for the low potential region. Note that the average distance to fertilizer seller is approximately 3 and 5 kilometers for the high and low potential regions respectively. Measures of the effect of accessibility to extension services and motorable roads show that households that are further away from these services have lower probability of purchasing fertilizers (Table 21). Each 10 kilometer away from a motorable road reduces fertilizer application rates per acre by half a kilogram in the low potential region.

Following theory, we use per kilogram relative nitrogen fertilizer-maize price ratio in our models. Hereafter we use prices to denote relative price of fertilizer to that of maize. The signs for the estimates on effect of prices on purchase and application rates per acre offer interesting results. Though prices have a negative but statistically insignificant influence on probability of purchase for all regions, higher prices significantly reduce fertilizer applied per acre for the low potential region. The conclusion to draw from this result is that those faced with higher prices will buy fertilizer but use it more sparingly i.e. less intensely or spread it thinly on the ground.

The size of land owned has no significant effect on amount of fertilizer applied per acre but has a positive influence on probability of buying in the higher potential region. Another measure of access to resources or wealth is the method of land preparation. There are three common means farmers use to prepare their land for planting, manual (hoe), ox-drawn ploughs, or tractor. The probability of purchasing fertilizer for households using ox or tractor is higher than those using manual hand-held tools. In addition these households use 3 kilos more on average per acre than those who use hand-held methods in the high potential region.

Based on land tenure the sample is divided into three categories, those who own land and have title to the land (and so can use it as collateral), those who own the land but have not yet received their title documents, and those that rent land. Households that rent land have higher

probability of purchasing fertilizer and use approximately 1 more kilogram of fertilizer per acre than those who own land but have no title documents in the high potential region. Probably renters could be intensifying fertilizer use so as to get good returns to their investment (knowing that the leasing is not permanent arrangement).

Households with heads who have more years of schooling have higher probability of purchasing fertilizer than those who do not (0.01 points per additional year of school) and also apply more fertilizer per acre (0.3 kilograms more per acre per year of schooling). The age of the head of household is significant and positive fertilizer use per acre but not significant for probability of purchasing fertilizer. Female headed households do not show any significant differences in demand from those headed by males.

Rainfall results show positive effect on probability of purchasing fertilizer and quantity applied per acre but this is not significant. However, for the low potential region moisture stress reduces application per acre. Clearly areas with poor rainfall trends are less likely to participate in fertilizer markets and for those who do, applications rates per acre are relatively less.

2.5 Conclusions and Recommendations

We test for and reject the assumption of exogeneity of distance to fertilizer seller and use appropriate IVs to mitigate possibility of biased estimates. Using IV techniques makes a contribution in providing estimates that are not biased. Tests of equivalence of estimates for the different models indicate that, if endogeneity is not controlled for, the estimates are smaller in magnitude. This has policy implications in that decisions that rely on such estimates will produce less than the desired effect.

This essay helps to explain why despite MVCR and AVCR being attractive (Essay 1), demand (purchase and intensification) for some areas is relatively low compared to other areas where fertilizer use is higher. There is heterogeneity in these estimates for the sample analyzed for this study. First, we observe that agro-ecological conditions are an important factor with some zones applying more fertilizer and more likely to participate in markets than other zones. Rainfall and moisture stress are important but with opposite effects on i) the location of fertilizer sellers relative to farms ii) households' fertilizer purchase decisions iii) quantity of fertilizer applied per acre.

Results show that distance to motorable road and extension service are correlated with the distance to where the nearest fertilizer seller can be found. Longer distances from farms to motorable roads or extension services are associated with relatively longer distances to nearest fertilizer sellers, indicating the effect of infrastructure (road and information) on private trade.

Households facing higher prices are further away from fertilizer sellers. Higher prices are associated with longer distances to sellers while less intensive fertilizer application practices are correlated with higher prices, with farmers applying less fertilizer per acre as prices rise in low potential region. There is a positive relationship between distances to fertilizer seller and

extension services. Extension services provide information on fertilizer use and this might influence demand for fertilizer and therefore leading retailers to open shop near extension offices.

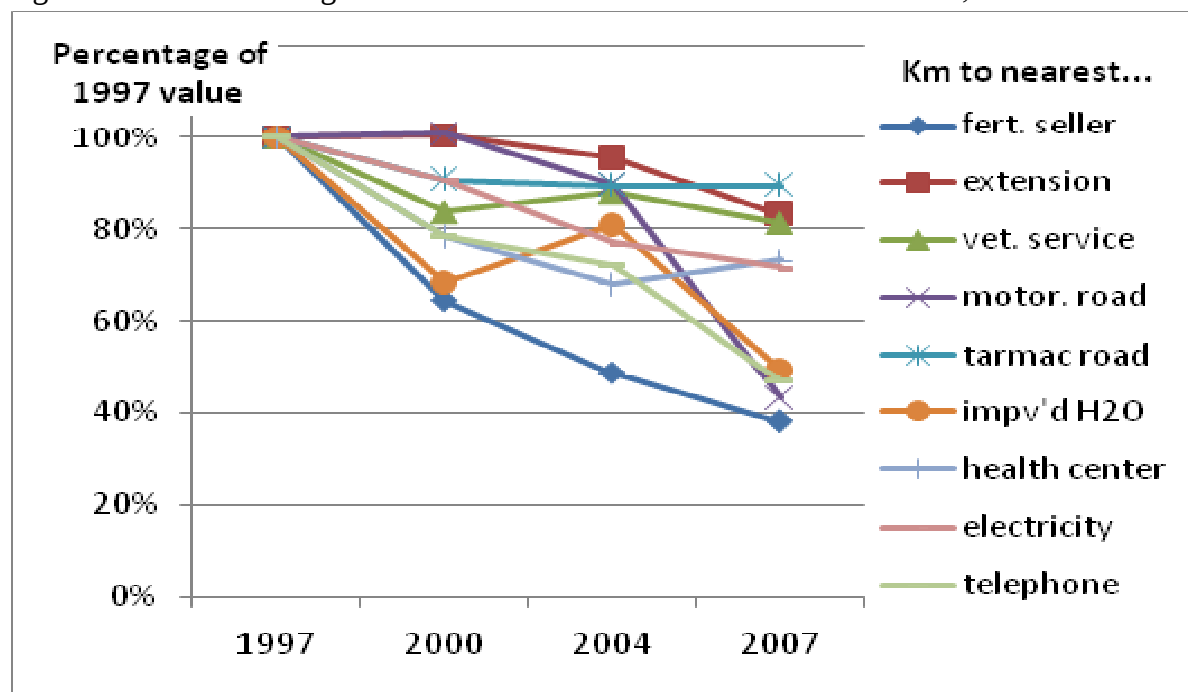
Households with more resources (assets, land) or access to technology like tractors and ox-plough for land preparation are also more likely to purchase fertilizer though applications rates per acre are not significantly different with those with fewer resources. This reveals that resource constraints need to be tackled in order to encourage producers to use fertilizer.

Households that own land and have title to their land and households that are renters have higher probability of purchasing fertilizer and also apply more fertilizer on the land than those who own land but do not yet have title to their land. Though we do not have sufficient data to infer whether those with titles were able to access credit we can only speculate at the value of having valid collateral documents. A clear-cut land policy with appropriate documents to title and property rights will be a catalyst for investments on farms and also create a market for rural assets which will enable efficient use of land by those who have the means and assets to do so.

APPENDIX

APPENDIX 2

Figure 11. Relative changes in indicators of access to markets and services, indexed to 1997



Source: Tegemeo Household Survey data 1997, 2000, 2004, 2007. The Constituency Development Fund (CDF), under which local authorities were given increased control of budget resources for local development, was established in 2003/04. All the 210 constituencies in Kenya are allocated 2.5 percent of the total government revenue for CDF funding. The sharp reduction in the distance to motorable roads and clean water between the 2004 and 2007 surveys is associated with this administrative reform, although causality cannot be inferred.

Table 11. Changes in Distance from Farm to the Nearest Fertilizer Seller

agro-ecological zones	1997	2000	2004	2007
Coastal Lowland	17.3	18	9.6	6.1
Eastern Lowland	8.1	4.7	3.7	2.7
Western Lowland	12.5	10.8	6.5	3.8
Western Transitional	6.1	4.4	2.8	2.7
High Potential	4.5	4.0	3.0	2.4
Western Highlands	2.8	2.1	1.4	1.2
Central Highlands	2.6	1.5	1.3	1.3

Note: Distance is measured in kilometers (1 kilometer=0.62 miles).

Table 12. Definition of Variables: Fertilizer Demand

Variable	Description of Variable	Units
Dependent:		
N	Quantity of N fertilizer applied per acre	kilograms
Prob	Binary variable (1=household used fertilizer,0=household did not use)	ordinal
Explanatory:		
DS	distance to nearest fertilizer seller	kilometers
DE	distance to extension service	kilometers
DR	distance to motorable road	kilometers
PR1	N-to-maize price ratio: kilo prices used	proportion
PR2	P-to-maize price ratio: kilo prices used	proportion
AS	value of household agricultural assets(in `0000)	Kenya Shillings
HA	land size	acres
AG	age of head of household	years
YL	number of years since household settled in this area	years
R	main season rainfall	millimeters
MS	proportion of 20-day periods during the growing season that had less than 40mm of rainfall	proportion
PN	N Fertilizer Price Per Kilo of Nutrient	Shillings
PP	P Fertilizer Price Per Kilo of Nutrient	Shillings

Note: To be consistent with Essay 1, Essay 2 uses nitrogen (N) fertilizer active ingredient as the dependent variable. Nitrogen (N) and Phosphorus (P) are very collinear due to the nature of fertilizer used in Kenya on maize. The most common fertilizer is DAP with a ratio 18:46:0 for N:P:K. Therefore, N and P are used in fairly fixed ratios and since the prices of N and P are derived from the “gross” fertilizer (DAP) price, there is a systematic relationship between these active ingredients. Table 25 and 26 indicate a correlation coefficient of 0.72 between N and P in the high potential region (0.83 for low potential region)

Table 13. High Potential Region: Descriptive Statistics for Continuous Variables

	Percentiles				
Variables	P10	P25	P50	P75	P90
N	0.00	3.27	9.00	18.00	30.36
P	0.00	2.01	6.54	10.03	15.04
DS	0.50	1.00	2.00	3.50	7.00
PR	6.16	7.12	8.33	9.87	11.90
AS	0.14	0.39	2.04	24.49	30.00
AG	37.00	44.00	54.00	64.00	72.00
HA	1.25	2.02	3.50	5.82	9.50
DE	1.00	2.00	3.00	6.00	10.00
DR	0.05	0.10	0.50	1.00	2.00
R	270.00	390.60	531.95	682.00	751.00
MS	0.00	0.00	0.17	0.38	0.62
ED	0.00	0.00	5.00	8.00	12.00
YL	12.00	18.00	28.00	35.00	44.00
PN	120.89	130.08	150.50	188.12	212.71
PP	47.30	50.90	58.89	73.61	83.23

Table 14. Low Potential Region: Descriptive Statistics for Continuous Variables

	Percentiles				
Variables	P10	P25	P50	P75	P90
N	0.00	0.00	0.00	0.00	2.20
P	0.00	0.00	0.00	0.00	0.94
DS	1.50	2.00	4.00	10.00	20.00
PR	6.42	7.68	8.73	10.24	12.31
AS	0.07	0.23	1.06	15.29	30.00
AG	38.00	46.00	56.00	65.00	73.00
HA	1.48	2.40	4.03	7.00	11.25
DE	1.00	2.00	4.60	8.40	13.00
DR	0.10	0.20	0.55	2.00	3.50
R	63.59	140.00	237.10	613.00	700.18
MS	0.00	0.00	0.22	0.69	1.00
ED	0.00	0.00	4.00	8.00	12.00
YL	7.00	15.50	27.00	37.00	47.00

Table 14 (cont'd)

PN	138.67	142.22	188.12	192.71	232.85
PP	54.26	55.65	73.61	75.41	91.12

Note: The variables are described or defined in Table 21

Table 15. Descriptive Statistics for Categorical Variables

	High Potential Region	Low Potential Region
Prob:		
Using Fertilizer	0.84	0.18
Not Using fertilizer	0.16	0.82
Land Preparation Technology (LP):		
Manual Implements	0.51	0.49
Ox-Plough	0.19	0.42
Tractor	0.30	0.09
Land Tenure (LT):		
Own Land and has title(LT1)	0.51	0.39
Own Land but without title (LT2)	0.39	0.56
Renting Land (LT3)	0.10	0.05
Asset Quartiles:		
Lowest Quartile	0.25	0.25
2 nd Quartile	0.25	0.25
3 rd Quartile	0.25	0.25
Highest Quartile	0.25	0.25
Land Size Quartiles:	0.25	0.25
Lowest Quartile	0.25	0.25
2 nd Quartile	0.25	0.25
3 rd Quartile	0.25	0.25
Highest Quartile	0.25	0.25
Gender of Head (G):		
Male head (FEM1)	0.76	0.63
Female, non-resident spouse (FEM2)	0.09	0.15
Female, no spouse (FEM3)	0.15	0.22
Agro-Zones (Zone):		
Coastal Lowland (1)		0.16
Eastern Lowlands (2)		0.41
Western Lowland (3)		0.43

Table 15 (cont'd)

Western Transitional (4)	0.17	
High Potential Maize Zone (5)	0.43	
Western Highlands (6)	0.16	
Central Highlands (7)	0.24	
Soils:		
Phaeozems & Luvisols (1)	0.13	0.05
Cambisols, Ferralsols, and Vertisols (2)	0.05	0.02
Regosols (3)	0.22	0.00
Rankers (4)	0.47	0.59
Other (5)	0.13	0.34

Note: Interpretation: in the high potential region 22% of the sample is covered by Regosol type of soils, 43% of the sample is under the high potential zone, and 51% of the sample plots are prepared using hand-held implements.

Table 16. High Potential Region: Correlation Matrix for Variables

	N	P	DS	PR	AS	AG	HA	DE	DR	R	MS	ED	YL	PN
N	1													
P	0.72*	1												
DS	-0.01	-0.03*	1											
PR	-0.05*	-0.03*	0.05*	1										
AS	0.04*	0.04*	0.16*	-0.03	1									
AG	-0.00	-0.01	-0.01	0.06*	-0.01	1								
HA	0.15*	0.07*	0.11*	-0.07*	0.31*	0.03*	1							
DE	-0.12*	-0.07*	0.16*	0.01	0.01	0.02	0.03	1						
DR	-0.05*	-0.01	0.07*	-0.08*	-0.02	-0.15*	0.01	0.15*	1					
R	0.07*	0.08*	-0.10*	-0.12*	-0.13*	0.02	0.03*	-0.02	0.07*	1				
MS	-0.04*	0.00	0.12*	-0.00	0.07*	-0.12*	-0.00	0.07*	0.06*	0.23*	1			
ED	0.19*	0.15*	-0.11*	-0.06*	-0.08*	-0.23*	0.08*	-0.05*	-0.00	0.30*	-0.28*	1		
YL	0.05*	0.07*	0.06*	0.05*	-0.02	-0.49*	0.00	-0.02	0.07*	-0.01	-0.00	0.18*	1	
PN	-0.17*	-0.11*	0.04*	0.05*	0.08*	-0.19*	-0.05*	0.01	0.07*	-0.31*	0.44*	-0.40*	-0.03*	1
PP	-0.17*	-0.11*	0.04*	0.05*	0.08*	-0.19*	-0.05*	0.01	0.07*	-0.31*	0.44*	-0.40*	-0.03*	1.00*

Note: significance is indicated at 5%. Note: These variables are described or defined in **Table 12**. Nitrogen (N), Phosphate (P), distance to fertilizer seller (DS), price ratio (PR), household assets (AS), age of household head (AG), size land owned (HA), distance to extension service (DE), distance to motorable road (DR), Rainfall (R), Moisture Stress (MS), education level of household head (ED), and number of years the household has lived in the location (YL), nitrogen price (PN), and phosphorous price (PP). The PN and PP have same coefficients because these nutrients occur in fixed proportions for Di-Ammonium Phosphate (DAP) fertilizer, the most common fertilizer used in maize production.

Table 17. Low Potential Region: Correlation Matrix for Variables

	N	P	DS	PR	AS	AG	HA	DE	DR	R	MS	ED	YL	PN
N	1													
P	0.83*	1												
DS	-0.19*	-0.18*	1											
PR	-0.10*	-0.04	-0.16*	1										
AS	-0.11*	-0.12*	0.06*	-0.03	1									
AG	-0.14*	-0.12*	-0.05*	0.09*	0.02	1								
HA	-0.02	-0.03	-0.00	0.06*	-0.14*	0.13*	1							
DE	-0.01	-0.00	0.33*	-0.12*	0.09*	-0.03	0.19*	1						
DR	0.04	0.00	0.21*	-0.10*	0.02	-0.10*	0.06*	0.33*	1					
R	-0.23*	-0.18*	0.07*	-0.03	0.29*	0.14*	-0.13*	0.07*	-0.02	1				
MS	0.28*	0.22*	-0.28*	0.14*	-0.25*	-0.01	0.19*	-0.01	-0.09*	-0.65*	1			
ED	0.23*	0.25*	-0.26*	0.02	-0.10*	-0.29*	0.04	-0.05*	-0.09*	0.03	0.27*	1		
YL	-0.02	-0.02	-0.03	0.05*	-0.06*	0.46*	0.15*	-0.03	-0.03	-0.02	0.03	-0.12*	1	
PN	-0.16*	-0.19*	0.32*	-0.09*	0.02	-0.11*	-0.13*	-0.02	0.08*	0.10*	-0.53*	-0.36*	-0.01	1
PP	-0.16*	-0.19*	0.32*	-0.09*	0.02	-0.11*	-0.13*	-0.02	0.08*	0.10*	-0.53*	-0.36*	-0.01	1.00*

Note: significance is indicated at 5%. Note: These variables are described or defined in Table 21. Nitrogen (N), Phosphate (P), distance to fertilizer seller (DS), price ratio (PR), household assets (AS), age of household head (AG), size land owned (HA), distance to extension service (DE), distance to motorable road (DR), Rainfall (R), Moisture Stress (MS), education level of household head (ED), and number of years the household has lived in the location (YL), nitrogen price (PN), and phosphorous price (PP). The PN and PP have same coefficients because these nutrients occur in fixed proportions for Di-Ammonium Phosphate (DAP) fertilizer, the most common fertilizer used in maize production.

Table 18. Gross Prices for Hybrid Seed, Grain, and DAP Fertilizer (Shs / kg)

agro-regional zones		1997	2000	2004	2007
Western Transitional	Seed	-	-	246	117
	Fertilizer	25	27	29	36
	Grain	11	14	14	13
High Potential	Seed	-	-	165	117
	Fertilizer	25	26	28	35
	Grain	11	12	13	11
Western Highlands	Seed	-	-	155	111
	Fertilizer	30	30	30	39
	Grain	12	13	15	14
Central Highlands	Seed	-	-	141	109
	Fertilizer	28	27	29	38
	Grain	13	14	14	13

Note: prices are not deflated or inflated using price index. “-“indicates data is missing.

Though the above table shows a general rise in prices, the real prices for both grain and inputs shown below indicate a decline.

Table 19. Indexed Prices for Hybrid Seed, Grain, and DAP Fertilizer (Shs / kg)

agro-regional zones		1997	2000	2004	2007
Western Transitional	Seed	-	-	198	117
	Fertilizer	55	48	39	36
	Grain	24	24	19	13
High Potential	Seed	-	-	224	117
	Fertilizer	55	45	38	35
	Grain	24	21	18	11
Western Highlands	Seed	-	-	209	111
	Fertilizer	66	54	41	39
	Grain	26	24	20	14
Central Highlands	Seed	-	-	191	109
	Fertilizer	60	48	39	38
	Grain	28	25	18	13

Note: Prices Indexed based on 2007 level. “-“indicates data is missing.

Table 20 . Factors Influencing Distance to Nearest Fertilizer Seller

	Pooled OLS		Fixed Effects		Random Effects	
VARIABLES	HP	LP	HP	LP	HP	LP
PR	0.03	-0.10*	0.08***	-0.11	0.04*	-0.25***
	(0.02)	(0.06)	(0.03)	(0.07)	(0.02)	(0.06)
DE	0.07***	0.17***	0.06***	-0.02	0.07***	0.13***
	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)	(0.02)
DR	0.12***	0.25***	-0.03	0.13*	0.07*	0.24***
	(0.03)	(0.07)	(0.04)	(0.07)	(0.03)	(0.07)
R	-0.00***	-0.01***	-0.00***	0.00	-0.00***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MS	1.67***	4.34***	1.42***	-2.86***	1.75***	3.35***
	(0.30)	(0.61)	(0.36)	(0.68)	(0.29)	(0.60)
AG	0.01**	-0.01	0.03**	0.03	0.01*	-0.02
	(0.00)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)
YL	-0.01***	-0.03***	-0.10***	-0.85***	-0.02***	-0.04***
	(0.00)	(0.01)	(0.02)	(0.05)	(0.01)	(0.01)
Asset :						
2 nd Quartile	-0.08	0.73*	-0.08	0.61	-0.04	0.99**
	(0.14)	(0.38)	(0.17)	(0.44)	(0.15)	(0.40)
3 rd Quartile	0.07	1.15***	-0.11	1.33***	0.07	1.45***
	(0.15)	(0.40)	(0.20)	(0.49)	(0.16)	(0.43)
Highest Quartile	0.40**	1.55***	-0.06	1.12*	0.39**	1.97***
	(0.18)	(0.47)	(0.24)	(0.60)	(0.19)	(0.51)
Land Size :						
2 nd Quartile	0.13	0.67*	0.02	0.38	0.11	0.68*
	(0.14)	(0.38)	(0.17)	(0.43)	(0.15)	(0.40)
3 rd Quartile	0.11	-0.12	0.14	-0.37	0.17	-0.11
	(0.15)	(0.39)	(0.20)	(0.47)	(0.16)	(0.41)
Highest Quartile	0.75***	0.78*	0.40*	0.03	0.71***	0.71
	(0.16)	(0.43)	(0.23)	(0.55)	(0.18)	(0.47)
Land Tenure (LT):						
LT2	0.09	0.05	0.04	0.25	0.05	0.06
	(0.12)	(0.29)	(0.15)	(0.31)	(0.12)	(0.29)

Table 20 (cont'd)

LT3	-0.34*	0.12	-0.20	0.77	-0.31*	0.13
	(0.18)	(0.64)	(0.20)	(0.63)	(0.18)	(0.63)
Gender (G):						
FEM2	-0.21	0.14	-0.50**	-0.61	-0.27	0.17
	(0.18)	(0.38)	(0.23)	(0.51)	(0.19)	(0.43)
FEM3	-0.20	-1.43***	-0.40	-0.18	-0.26	-1.73***
	(0.14)	(0.34)	(0.30)	(0.57)	(0.17)	(0.39)
Zone:						
Zone2		-6.94***				-7.32***
		(0.46)				(0.58)
Zone3		-2.93***				-2.74***
		(0.62)				(0.71)
Zone5	-0.52***					
	(0.20)					
Zone6	-2.07***					
	(0.18)					
Zone7	-1.61***					
	(0.25)					
Soils:						
Soil2	-1.82***	5.90***			-1.88***	6.53***
	(0.25)	(1.11)			(0.33)	(1.44)
Soil3	-1.00***				-1.04***	
	(0.19)				(0.24)	
Soil4	-0.32*	0.11			-0.32	-0.06
	(0.17)	(0.65)			(0.22)	(0.89)
Soil5	1.34***	3.25***			1.26***	3.00***
	(0.28)	(0.78)			(0.36)	(1.04)

Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Note: HP=High Potential; LP=Low Potential; for categorical variables the group that is ranked first or quartile is dropped (so compare estimates for the other groups based on this dropped group). For land and assets we drop the lowest rank. Gender (male dummy is dropped); for land tenure we drop “own with title”; Zone (coastal and western transitional zone are dropped). The abbreviated variable names are explained in detail in Table 12 and Table 15.

Table 21 . Average Marginal Effects for Factors Influencing the Probability of Purchasing Fertilizer

	Pooled		CRE	
VARIABLES	HP	LP	HP	LP
DS	-0.01*	-0.02***	-0.00	-0.02***
	(0.00)	(0.01)	(0.00)	(0.01)
PR	-0.00	-0.00	-0.01	-0.00
	(0.00)	(0.00)	(0.18)	(0.33)
AS	-0.00***	0.01**	-0.00	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
AG	0.00	0.00	0.00	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
ED	0.01***	0.01***	0.00**	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
HA	0.00	0.00	0.01***	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
DE	-0.00**	0.01***	-0.00	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
DR	-0.05***	0.02*	-0.01*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
R	0.00**	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
MS	0.08	-0.11	-0.03	-0.06
	(0.06)	(0.09)	(0.05)	(0.09)
G:				
FEM2	0.03	0.02	0.01	-0.04
	(0.02)	(0.02)	(0.02)	(0.03)
FEM3	-0.01	-0.02	0.01	-0.04
	(0.02)	(0.02)	(0.02)	(0.04)
LP:				
Ox	0.07***	-0.07***	0.07***	-0.00
	(0.02)	(0.02)	(0.02)	(0.03)
Tractor	0.15***	-0.07	0.15***	-0.04
	(0.02)	(0.05)	(0.02)	(0.05)

Table 21 (cont'd)

LT:				
Own Land	-0.04***	0.07***	-0.02	0.05***
	(0.01)	(0.02)	(0.01)	(0.02)
Rent	0.07***	0.12***	0.05***	0.08*
	(0.02)	(0.05)	(0.02)	(0.04)
Zone:				
Zone2		0.18***		0.01
		(0.03)		(0.06)
Zone3		0.09**		0.41***
		(0.05)		(0.06)
Zone5	0.18***		0.41***	
	(0.03)		(0.04)	
Zone6	0.26***		0.32***	
	(0.03)		(0.05)	
Zone7	0.27***		0.52***	
	(0.04)		(0.04)	
Soils:				
Soil2	0.13***		-0.02	-0.11*
	(0.03)		(0.05)	(0.06)
Soil3	0.07***		0.04	
	(0.02)		(0.03)	
Soil4	0.08***	0.13***	0.04	0.06
	(0.02)	(0.03)	(0.03)	(0.06)
Soil5	-0.07*	0.07*	-0.04	0.12
	(0.04)	(0.04)	(0.05)	(0.07)
M-C:				
DS			-0.00	-0.01***
			(0.00)	(0.01)
PR			0.01	0.01
			(0.18)	(0.33)
AS			-0.00***	-0.00
			(0.00)	(0.00)
AG			-0.00	-0.00
			(0.00)	(0.00)

Table 21 (cont'd)

ED			0.00	-0.00
			(0.00)	(0.01)
HA			-0.00	-0.00
			(0.00)	(0.00)
DE			-0.00	0.01*
			(0.00)	(0.00)
DR			-0.04***	0.00
			(0.01)	(0.01)
R			0.00**	-0.00**
			(0.00)	(0.00)
MS			-0.60***	0.06
			(0.17)	(0.35)

These estimates are derived from the table below which contains the raw regression results. The Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1: Note: These are results from OLS on pooled data and Correlated Random Effects (CRE) probit model results for High and Low Potential Regions (HP & LP). The CRE approach incorporates the M-C device. Gender (male dummy is dropped);soil (category for “phaeozems & Luvisols” is dropped), Zone (western transitional zone in HP and coastal lowland zone in LP region are dropped), land preparation (manual category is dropped),and tenure category of “own land with title” is dropped. The coefficients for the existing categories for these variables relate to the dropped ones. abbreviated variable names are explained in detail in Table 12 and Table 15.

Table 22. Probit Regression for Factors Influencing the Probability of Purchasing Fertilizer

VARIABLES	Pooled		CRE	
	HP	LP	HP	LP
DS	-0.03	-0.18***	-0.03	-0.25***
	(0.02)	(0.05)	(0.04)	(0.07)
DS^2	0.00	0.01***	0.00	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
PR	-0.14**	-0.08	-0.17	-0.01
	(0.06)	(0.07)	(1.39)	(2.52)
PR^2	0.01**	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
AS	-0.04**	0.08**	-0.04*	0.08
	(0.02)	(0.03)	(0.02)	(0.05)

Table 22 (cont'd)

AS^2	0.00*	-0.00***	0.00**	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
AG	0.03**	-0.01	0.02	0.02
	(0.01)	(0.02)	(0.02)	(0.04)
AG^2	-0.00**	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
ED	0.03	0.02	0.04	0.04
	(0.02)	(0.03)	(0.03)	(0.05)
ED^2	0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
HA	0.01	0.03	0.05***	0.09*
	(0.01)	(0.03)	(0.02)	(0.05)
HA^2	0.00	-0.00	-0.00	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
DE	-0.02*	0.09***	-0.01	0.08*
	(0.01)	(0.03)	(0.02)	(0.04)
DE^2	0.00	-0.00***	0.00	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
DR	-0.30***	0.19*	-0.13*	0.20
	(0.05)	(0.10)	(0.07)	(0.14)
DR^2	0.03***	-0.04**	0.01*	-0.04*
	(0.01)	(0.02)	(0.01)	(0.02)
R	-0.00**	0.00	-0.00*	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.00***	-0.00	0.00*	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
MS	3.48***	0.64	2.92**	0.08
	(0.91)	(1.02)	(1.24)	(1.40)
MS^2	-2.01***	-1.34*	-2.02**	-0.63
	(0.70)	(0.75)	(0.95)	(1.03)
G:				
FEM2	0.14	0.10	0.06	-0.34
	(0.11)	(0.13)	(0.18)	(0.25)
FEM3	-0.03	-0.11	0.08	-0.30
	(0.08)	(0.14)	(0.19)	(0.30)

Table 22 (cont'd)

LP:				
Ox	0.30***	-0.37***	0.44***	-0.03
	(0.09)	(0.13)	(0.15)	(0.22)
Tractor	0.85***	-0.40	1.31***	-0.30
	(0.11)	(0.29)	(0.17)	(0.39)
LT:				
Own Land	-0.17***	0.37***	-0.16*	0.40***
	(0.06)	(0.11)	(0.09)	(0.15)
Rent	0.44***	0.66***	0.52***	0.58*
	(0.13)	(0.23)	(0.18)	(0.31)
Zone:				
Zone2		1.10***		0.07
		(0.21)		(0.52)
Zone3		0.69**		4.72***
		(0.34)		(1.49)
Zone5	0.67***		2.05***	
	(0.11)		(0.25)	
Zone6	1.12***		1.47***	
	(0.12)		(0.22)	
Zone7	1.19***		3.36***	
	(0.16)		(0.39)	
Soils:				
Soil2	0.70***		-0.15	-6.79
	(0.17)		(0.32)	(254.40)
Soil3	0.32***		0.33	
	(0.11)		(0.22)	
Soil4	0.38***	0.88***	0.27	0.47
	(0.10)	(0.25)	(0.19)	(0.54)
Soil5	-0.27*	0.53*	-0.27	0.90
	(0.15)	(0.31)	(0.29)	(0.60)
M-C:				
DS			-0.03	-0.11**
			(0.03)	(0.04)
PR			0.10	0.06
			(1.39)	(2.52)

Table 22 (cont'd)

AS			-0.03***	-0.00
			(0.01)	(0.02)
AG			-0.01	-0.02
			(0.01)	(0.02)
ED			0.01	-0.01
			(0.02)	(0.04)
HA			-0.00	-0.02
			(0.02)	(0.04)
DE			-0.02	0.06*
			(0.02)	(0.03)
DR			-0.28***	0.04
			(0.07)	(0.10)
R			0.00**	-0.01**
			(0.00)	(0.01)
MS			-4.66***	0.48
			(1.30)	(2.70)
Constant	-1.34**	-2.33***	-1.78	-0.55
	(0.55)	(0.85)	(1.44)	(2.78)

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1: Note: The results from these regressions are given in their raw form (this table) and also as marginal effects (in table above). These are results from OLS on pooled data and Correlated Random Effects (CRE probit model results for High and Low Potential Regions (HP & LP). The CRE approach incorporates the M-C device. Gender (male dummy is dropped);soil (category for “phaeozems & Luvisols” is dropped), Zone (western transitional zone in HP and coastal lowland zone in LP region are dropped), land preparation (manual category is dropped),and tenure category of “own land with title” is dropped. The interpretation of the existing categories for these variables relate to the dropped ones. The abbreviated variable names are explained in detail in Table 12 and Table 15.

Table 23. Marginal Effects for Factors Influencing Quantity Applied

VARIABLES	High Potential			Low Potential		
	IV	POLS	CRE	IV	POLS	CRE
DS	3.46	0.22*	0.11	-0.21	-0.04**	-0.04**
	(2.94)	(0.12)	(0.12)	(0.21)	(0.02)	(0.02)
DE	-0.46	-0.09	-0.08	0.06	0.04*	0.02
	(0.34)	(0.08)	(0.07)	(0.05)	(0.02)	(0.02)
DR	-1.12***	-0.89***	-0.75***	0.15*	0.06	0.07*

Table 23 (cont'd)

	(0.34)	(0.30)	(0.28)	(0.09)	(0.05)	(0.04)
PR	-0.08	-0.06	-0.04	-0.20*	-0.05	-0.02
	(0.13)	(0.12)	(0.10)	(0.12)	(0.04)	(0.03)
AS	-0.04	-0.02	-0.05	0.02	0.02	0.01
	(0.05)	(0.05)	(0.04)	(0.02)	(0.02)	(0.02)
AG	0.17***	0.17***	0.15***	-0.03	0.00	0.00
	(0.05)	(0.04)	(0.04)	(0.02)	(0.01)	(0.01)
ED	0.28***	0.26***	0.25***	-0.02	0.04	0.05**
	(0.08)	(0.08)	(0.07)	(0.05)	(0.02)	(0.02)
HA	0.01	0.03	-0.02	-0.00	-0.02	-0.02
	(0.08)	(0.07)	(0.07)	(0.03)	(0.02)	(0.02)
R	0.00*	0.00	0.00	0.00*	0.00	0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MS	-6.93	2.10	-0.02	-0.99**	-1.23***	-0.93***
	(3.40)	(2.10)	(1.94)	(0.46)	(0.38)	(0.35)
G:						
FEM2	-0.69	-0.79	-0.90	-0.27	-0.26	-0.27
	(0.91)	(0.88)	(0.82)	(0.23)	(0.20)	(0.19)
FEM3	1.53	0.99	0.37	-0.96*	-0.29	-0.17
	(1.11)	(1.04)	(1.00)	(0.56)	(0.22)	(0.21)
LP:						
Ox	2.19*	2.31***	3.08***	-0.28	-0.42***	-0.27*
	(1.19)	(0.71)	(0.73)	(0.18)	(0.15)	(0.15)
Tractor	6.33***	7.08***	6.65***	-0.29	-0.25	-0.25
	(1.14)	(0.72)	(0.74)	(0.26)	(0.22)	(0.23)
LT:						
Own Land	-1.19**	-1.11**	-0.89*	0.33**	0.37***	0.26**
	(0.48)	(0.46)	(0.47)	(0.13)	(0.11)	(0.11)
Rent	0.84	0.47	-0.08	0.14	0.23	0.17
	(0.75)	(0.71)	(0.68)	(0.29)	(0.25)	(0.24)
Zone:						
Zone2				-0.32	1.39***	1.51***
				(1.31)	(0.22)	(0.27)
Zone3				-0.83	0.65**	0.83**
				(1.12)	(0.30)	(0.34)

Table 23 (cont'd)

Zone5	2.47***	1.98**	3.39***			
	(0.88)	(0.83)	(1.30)			
Zone6	1.18	-1.05	1.07			
	(1.37)	(0.74)	(1.24)			
Zone7	8.68***	7.58***	8.30***			
	(1.15)	(1.07)	(1.63)			
Soils:						
Soil2	4.71***	2.87***	2.70*	0.11	-0.25	-0.00
	(1.39)	(1.03)	(1.42)	(0.60)	(0.48)	(0.64)
Soil3	-3.68***	-4.85***	-5.16***			
	(0.91)	(0.75)	(1.03)			
Soil4	0.04	-0.24	-0.77	1.08***	1.27***	1.41***
	(0.70)	(0.67)	(0.93)	(0.32)	(0.27)	(0.37)
Soil5	-10.27***	-9.08***	-9.33***	1.70***	0.98***	0.95**
	(1.38)	(1.13)	(1.52)	(0.64)	(0.32)	(0.44)
M-C:						
ED	-0.00	-0.06	0.01	0.03	0.00	-0.00
	(0.10)	(0.10)	(0.11)	(0.04)	(0.03)	(0.03)
PR	-0.70***	-0.82***	-0.90***	-0.03	-0.25***	-0.26***
	(0.19)	(0.19)	(0.23)	(0.18)	(0.05)	(0.06)
AS	-0.01	-0.01	-0.02	0.01	0.00	0.00
	(0.04)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)
AG	-0.18***	-0.19***	-0.16***	0.01	-0.01	-0.01
	(0.05)	(0.05)	(0.05)	(0.02)	(0.01)	(0.01)
HA	0.13*	0.18***	0.21***	-0.02	-0.02	-0.01
	(0.07)	(0.06)	(0.07)	(0.02)	(0.02)	(0.02)
DE	-0.59***	-0.60***	-0.60***	0.08	0.02	-0.00
	(0.10)	(0.10)	(0.12)	(0.05)	(0.02)	(0.02)
DR	0.07	0.35	0.32	0.01	-0.00	0.01
	(0.34)	(0.30)	(0.37)	(0.06)	(0.05)	(0.06)

Note: IV=Instrumental Variable; POLS=Pooled OLS; CRE=Correlated Random Effects

Table 24. Regression Estimates for Factors Influencing Quantity Applied

	High Potential			Low Potential		
VARIABLES	IV	POLS	CRE	IV	POLS	CRE
DS	3.94	0.01	-0.04	-0.25	-0.05	-0.09*
	(3.01)	(0.18)	(0.18)	(0.24)	(0.05)	(0.04)
DS^2	-0.01	-0.00	-0.00	0.03***	0.00	0.00**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
DE	-0.65	-0.11	-0.10	0.10	0.05*	0.03
	(0.44)	(0.09)	(0.08)	(0.08)	(0.03)	(0.02)
DE^2	0.01	0.00	0.00	-0.00	-0.00*	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
DR	-1.80*	-0.73**	-0.74**	0.14	0.05	0.07
	(0.96)	(0.37)	(0.35)	(0.10)	(0.05)	(0.05)
DR^2	0.24*	0.09**	0.09**	0.00	0.00	-0.00
	(0.13)	(0.04)	(0.04)	(0.00)	(0.00)	(0.00)
PR	1.11	0.26	0.36	-0.33	-0.06	-0.01
	(0.79)	(0.36)	(0.34)	(0.32)	(0.08)	(0.07)
PR^2	-0.06	-0.01	-0.02	0.01	0.00	0.00
	(0.04)	(0.02)	(0.02)	(0.01)	(0.00)	(0.00)
AS	-0.13	0.14	-0.05	0.06	0.05	0.02
	(0.20)	(0.12)	(0.12)	(0.05)	(0.04)	(0.04)
AS^2	0.00	-0.00	0.00	-0.00	-0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
AG	0.28*	0.25**	0.29**	-0.06	-0.04	-0.04
	(0.16)	(0.11)	(0.12)	(0.04)	(0.03)	(0.03)
AG^2	-0.00	-0.00	-0.00	0.00	0.00*	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ED	0.30	0.07	0.18	-0.05	0.04	0.06*
	(0.21)	(0.13)	(0.12)	(0.11)	(0.04)	(0.04)
ED^2	-0.00	0.01*	0.00	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
HA	-0.11	0.03	-0.02	-0.03	-0.02	-0.02
	(0.13)	(0.07)	(0.07)	(0.04)	(0.03)	(0.03)
HA^2	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 24 (cont'd)

R	-0.00	-0.00	-0.01	0.01**	0.00*	0.00**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
R^2	0.00	0.00	0.00	-0.00**	-0.00*	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MS	10.86	18.86***	14.84**	-6.18*	-2.95**	-3.23***
	(10.06)	(6.73)	(6.19)	(3.44)	(1.16)	(1.08)
MS^2	-14.15**	-14.14***	-11.09**	4.33	1.53*	1.95**
	(7.01)	(5.14)	(4.74)	(2.86)	(0.86)	(0.80)
G:						
FEM2	0.31	-0.69	-0.80	-0.26	-0.26	-0.27
	(1.41)	(0.86)	(0.81)	(0.22)	(0.20)	(0.19)
FEM3	2.20	0.32	0.01	-0.93	-0.27	-0.15
	(2.10)	(1.02)	(0.99)	(0.82)	(0.22)	(0.21)
LP:						
Ox	2.19*	2.31***	3.08***	-0.28	-0.42***	-0.27*
	(1.19)	(0.71)	(0.73)	(0.18)	(0.15)	(0.15)
Tractor	6.33***	7.08***	6.65***	-0.29	-0.25	-0.25
	(1.14)	(0.72)	(0.74)	(0.26)	(0.22)	(0.23)
LT:						
Own Land	-1.22*	-1.22***	-0.94**	0.24*	0.35***	0.25**
	(0.68)	(0.45)	(0.46)	(0.14)	(0.11)	(0.11)
Rent	1.01	0.20	-0.34	0.13	0.23	0.16
	(1.34)	(0.70)	(0.67)	(0.29)	(0.25)	(0.24)
Zone:						
Zone2				-0.23	1.16***	1.05***
				(1.72)	(0.31)	(0.40)
Zone3				2.25**	2.33***	2.76***
				(1.08)	(0.79)	(1.01)
Zone5	8.12***	6.39***	22.24***			0.00
	(2.09)	(0.92)	(7.05)			(0.00)
Zone6	6.93	-0.11	15.65**			0.00
	(5.98)	(0.85)	(6.82)			(0.00)
Zone7	21.84***	17.14***	32.19***			0.00
	(5.00)	(1.42)	(6.90)			(0.00)

Table 24 (cont'd)

Soils:						
Soil2	6.36	0.21	0.02	-0.80	-0.84*	-0.53
	(5.55)	(1.07)	(1.50)	(0.75)	(0.51)	(0.68)
Soil3	-0.14	-4.10***	-4.12***			0.00
	(3.61)	(0.77)	(1.06)			(0.00)
Soil4	-0.08	-0.70	-1.13	0.51	0.82***	0.92**
	(1.37)	(0.66)	(0.92)	(0.59)	(0.31)	(0.43)
Soil5	-12.70***	-8.90***	-9.12***	1.60	1.00***	0.99**
	(3.72)	(1.13)	(1.52)	(1.03)	(0.35)	(0.48)
M-C:						
DS		0.36**	0.42**		-0.06***	-0.07**
		(0.14)	(0.17)		(0.02)	(0.03)
DE	-0.43***	-0.50***	-0.50***	0.09	0.04*	0.02
	(0.16)	(0.10)	(0.12)	(0.09)	(0.02)	(0.02)
DR	-1.19	-0.10	-0.19	0.01	0.01	0.02
	(0.98)	(0.30)	(0.37)	(0.07)	(0.05)	(0.06)
PR	-0.04	-0.57***	-0.59**	0.01	-0.21***	-0.21***
	(0.54)	(0.19)	(0.23)	(0.26)	(0.05)	(0.06)
AS	-0.06	-0.02	-0.03	0.01	0.00	-0.00
	(0.06)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)
AG	-0.13	-0.19***	-0.17***	0.01	-0.01	-0.02
	(0.08)	(0.05)	(0.05)	(0.03)	(0.01)	(0.01)
ED	0.10	-0.07	-0.02	0.05	-0.00	-0.01
	(0.19)	(0.10)	(0.11)	(0.06)	(0.03)	(0.03)
HA	0.02	0.17***	0.20***	0.02	0.01	0.02
	(0.17)	(0.06)	(0.07)	(0.03)	(0.02)	(0.03)
R	-0.01	-0.01	-0.01	-0.01	-0.01***	-0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
MS	-40.84***	-38.04***	-35.70***	-3.88	-2.25	-1.66
	(8.92)	(4.83)	(6.32)	(4.36)	(1.61)	(2.11)
Constant	-3.37	15.90***	0.00	9.73	6.34***	5.87***
	(18.43)	(5.78)	(0.00)	(6.07)	(1.57)	(1.99)

Note: IV=Instrumental Variable; POLS=Pooled OLS; CRE=Correlated Random Effects

BIBLIOGRAPHY

BIBLIOGRAPHY 2

- Adesina, A., and Chianu, J. (2002). Determinants of farmers' adoption and adaptation of alley farming technology in Nigeria. *Agroforestry Systems* 55: 99–112, 2002.
- Allgood, J. H., and Kilungu, J. (1996). An Appraisal of the Fertilizer Market in Kenya and Recommendations for Improving Fertilizer Use Practices by Smallholder Farmers: A Field Report. Muscle Shoals, AL: International Fertilizer Development Center (IFDC).
- Barrett, C. B., Bell, R., Erin C. Lentz, E. C., and Maxwell, D. G. (2009). Market information and food insecurity response analysis. *Food Security*, Springer Netherlands. 1: 151-168.
- Burke, W.J. 2009. Fitting and interpreting Cragg's tobit alternative using Stata. *Stata Journal* , StataCorp LP, vol. 9(4), pages 584-592, December
- Blaylock, J., Smallwood, D., and Blisard, N. (1992). How Did Household Characteristics Affect Food Spending in 1980-88?, Economic Research Service, U.S. Department of Agriculture, Agriculture Information Bulletin No. 643, 19 pp. February.
- Chamberlain, G. (1984). Panel data, Chapter 22 in Z. Giliches and M. Intrilligator,, eds., *Handbook of Econometrics*, North-Holland, Amsterdam: 1247 - 1318.
- Crawford, E. W., Jayne, T. S., and Kelly, V. A. (2006). Alternative Approaches for Promoting Fertilizer Use in Africa. *Agriculture and Rural Development Discussion Paper 22*. Washington, D.C.: World Bank.
- Cragg, J. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica* 39:829–844.
- Croppenstedt, A., and Muller, C. (2000). The Impact of farmers health and nutritional status on their productivity and efficiency evidence from Ethiopia. *Economic Development and Cultural Change*, 48(3), 475–502.
- Diagne, A., and Zeller, M. (2001). Access To Credit And Its Impact On Welfare Malawi. *International Food Policy Research Institute* Washington, D.C.
- Dorward, A., Fan, S., Kydd, J., Lofgren, H., Morrison, J., Poulton, C., Rao, N., Smith, L., Tchale, H., Thorat, S., Urey, I., and Wobst, P. (2004). "Rethinking Agricultural Policies for Pro-Poor Growth." *Natural Resources Perspectives* 94. London: ODI, September.
- Duflo, E., Kremer, M., and Robinson, J. (2007). *Understanding Technology Adoption*:

- Fertilizer in Western Kenya, Preliminary Results from Field Experiments.
Mimeo, MIT
- FAO. (2000). The Challenges of Sustainable Forestry Development in Africa. Twenty-first FAO Regional Conference for Africa. Yaounde, FAO.
- Gregory, I. (2006). The role of input vouchers in pro poor growth. IFDC Background Paper, Africa Fertilizer Summit, 9-13 June 2006. Abuja, Nigeria.
- Goetz, S. (1992). "A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa." *American Journal of Agricultural Economics* 74:444–452.
- Haines, P. S., Guilkey, D.K., Hungerford, D.W., and Popkin, B.M.. (1988). Eating out: who and where. In: *Outlook 89*. Washington, DC: U.S. Department of Agriculture.
- Hausman, J., Hall, B. H., and Griliches, Z. (1984). Econometric Models for Count data with an Application to the Patent-R&D Relationship. *Econometrica*, Econometric Society, vol. 52(4), pages 909-38.
- Heckman, J. J. (1981), "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden. Cambridge, MA: MIT Press, 179–195.
- IFDC. 2001. *An assessment of fertilizer prices in Kenya and Uganda: Domestic prices vis-à-vis international market prices*. International Center for Soil Fertility and Agricultural Development (IFDC), Muscle Shoals, Alabama, U.S.A. Photocopy.
- Jayne, T. S., and Argwings-Kodhek, G. (1997). Consumer response to maize market liberalization. *Food Policy* 22 (5): 447-458.
- Kelly, V. A. 2006. Factors Affecting Demand for Fertilizer in Sub-Saharan Africa: Agriculture and Rural Development Discussion Paper 23. World Bank, Washington, DC.
- Key, N., Sadoulet, E., and de Janvry, A. (2000). "Transactions costs and agricultural household supply response." *American Journal of Agricultural Economics* 82:245–259.
- Kherallah, M., Delgado, C., Gabre-Madhin, E., Minot, N., and Johnson, M. (2002). *Reforming Agricultural Markets in Africa*. Baltimore, MD: IFPRI/Johns Hopkins University Press.
- Ministry of Agriculture. (2008). Economic review of agriculture: 2008. Central Planning and Project Monitoring Unit. Nairobi, Kenya: Ministry of Agriculture.

- Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica* 46(1): 69-85.
- Morris, M., Kelly, V. A., Kopicki, R., and Byerlee, D. (2007). Promoting increased fertilizer use Africa: Lessons learned and good practice guidelines. Washington, D.C.: World Bank.
- Omamo, S. W., Williams, J. C., Obare, G. A., and Ndiwa, N. N. (2002). Soil fertility management on small farms in Africa: evidence from Nakuru District, Kenya. *Food Policy* 27(2): 159-170.
- Omamo, S.W., and Mose, L. O. (2001). "Fertilizer Trade under Market Liberalization: Evidence from Kenya." *Food Policy* 26(1): 1–10.
- Poulton, C., Kydd, J., and Dorward, A. (2006). Overcoming Market Constraints on Pro-Poor Agricultural Growth in Sub-Saharan Africa." *Development Policy Review* 24, no. 3(2006a): 243-277.
- Scherr, S. (1999). Soil Degradation: A Threat to Developing-country Food Security by 2020? Food, Agriculture, and the Environment Discussion Paper 27. Washington, DC: International Food Policy Research Institute.
- Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel data*. Cambridge, MA: MIT Press, 2002.
- Zhao, I. (2009). China Fertilizer Trade: Situation and Outlook. British Sulphur Consultants (Beijing)