AN ANALYSIS OF STOCHASTIC MAIZE PRODUCTION FUNCTIONS IN KENYA

by

ASHLEY D. JONES

B.S., University of Arkansas, 2009

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2011

Approved by:

Major Professor
Dr. Timothy J. Dalton

Copyright

Abstract

In Kenya, agriculture governs the country's fiscal economy, and this reliance on agriculture can cause both economic and hunger problems, a result of the country's dependence upon rainfall for agricultural production. Kenyans must find ways to combat severe drought conditions; this can be accomplished through the adoption of inputs that decrease the probability of crop failure. The objective of this research is to determine whether variability exists in Kenyan maize yields, and whether or not specific inputs, specifically hybrid varieties, are either variance/skewness increasing or decreasing.

The data used for this study was collected from a survey, designed by Egerton University's Tegemeo Institute of Agricultural Policy and Development and Michigan State University, and administered in Kenya in the following years: 1997, 2000, 2004, and 2007. The survey identified factors of crop and field level production, such as inputs, crop mix, marketing data, and demographic information. This research makes use of only the 2007 data, comprising 1,397 households in total. The objectives of this thesis aim to go beyond the scope of typical production function regressions where yield is a function of a set of inputs, by examining further moments of yield, variance, and skewness to determine whether variability exists in Kenyan maize yields.

Results indicate that variability does exist within Kenyan maize yields, often a result of differing input levels among households. In terms of overall impact of each variable on mean, variance, and skewness of maize yields, seed quantity, nitrogen use, and hybrid seed contribute the most to influencing these factors. In contrast, years of experience with hybrid maize, land tenure, terraced land and labor have the least influence on mean, variance and skewness within this research. Results also bring to light the popular debate against hybrid varieties versus open pollinated (OPV) or traditional varieties, and identify hybrid varieties as a source of variability in mean, variance and skewness of yields. Hybrid varieties should be paired with the knowledge of how to maximize yield in conjunction with other inputs, to give Kenya the opportunity to see substantial productivity gains throughout the country, especially in arid and semi-arid regions.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to thank my major professor, Dr. Timothy Dalton, for his endless support and patience during this research. His expertise in production economics and knowledge of African production systems was essential to this research. I would also like to thank my committee members, Dr. Jeff Williams and Dr. Jason Bergtold for their guidance and support throughout this research.

Second, I would like to thank Egerton University's Tegemeo Institute of Agricultural Policy and Development, as well as Michigan State University for providing a data set, specifically, one that was well organized. Their efforts made this research easier to manage and my research goals obtainable. Their efforts are greatly appreciated.

Next, I would like to thank my colleagues on the fourth floor for their infinite support in the process of completing my thesis. This experience has been both rewarding and enjoyable because of your influence in this year long journey.

I would also like to thank the Department of Agricultural Economics and the Graduate School at Kansas State University for providing funding throughout my graduate school career.

Last and certainly not least, I would like to thank my family and friends for their support in achieving my educational goals. Their encouragement, advice, and well wishes were essential to successfully completing the Agricultural Economics graduate program.

I am forever thankful for the individuals mentioned above; they are the reason I have completed my thesis requirement. As a result, this experience has been one that has taught me countless things and is one that will be a learning experience for all things to come in the future.

Chapter 1 - Introduction

Kenya is a country located on the east coast of Africa, bordered by Somalia, Ethiopia, Sudan, Uganda, and Tanzania. While this country is classified as one of the poorest developing nations in the world, it has advanced over the last few decades, particularly in its political, economic, and socioeconomic sectors. Agriculture governs Kenya's fiscal economy, resulting in many challenges due to the country's reliance on rainfall. The country is hindered by droughts that reduce agricultural output throughout the country. Without adequate irrigation and variable rainfall, households must find other ways to minimize their risk in order to achieve maximum output levels; especially with agricultural production split between subsistence farming and marketable agriculture.

1.1 Background

Looking at maize production in Kenya, the past few years have seen an increase in total production levels. Agriculture in Kenya is made up largely of maize production, as it is easily marketed. Maize also makes up a large portion of a Kenyan's daily caloric intake and is the largest staple food item among individuals (Karanja, 2003; Nyoro, 2002). Approximately three-fourths of all maize production is from small-scale farmers, who produce for both subsistence needs and marketing opportunities (Nyoro, 2002). Globally, maize yields average 3.5 tons per hectare in developing countries (Edmeades, 2008). In 1993, maize production in southern Africa totaled 23.5 million tons, an increase from 1992 levels of 12.5 million tons during a drought period (Edmeades, 2008). Looking specifically at Kenya, the country produced 2,367,237 metric tons of maize in 2008 (Food and Agriculture Organization, 2010).

Despite an increase in production levels, shortages exist both globally and locally. These shortages are often a result of stochastic factors outside of the farmer's control, such as droughts, pests, or lack of access to inputs. Such shortages in Kenya totaled 180,000 to 540,000 metric tons in the 2002 and 2003 growing season (Nyoro, Kirimi, and Jayne, 2004). In particular, drought loss will likely increase to 9.1 million metric tons annually, a result of increasing susceptibility to drought in Kenya (Edmeades, 2008). In terms of drought loss, fifteen percent of total maize grown in the world is lost as a result of drought (Edmeades, 2008). Rainfed maize production makes up the majority of the world's 160 million hectares of maize produced (Edmeades, 2008).

In the case of Africa, a large portion of the continent relies on rainfall to meet agricultural production needs. As a result of dependence on rainfall, many regions of the country must deal with drought in the same time period, creating regional shortages (Edmeades, 2008).

Specifically, Eastern Kenya can benefit significantly from improved seed varieties, aimed at both improving average yields, as well as stabilizing yields over time, because of strenuous climate conditions in the region. While climate varies within different regions of the country, Eastern Kenya suffers frequently from inadequate rainfall. The coastal areas of the country have a tropical climate, while the inner regions are arid and prone to drought conditions (Library of Congress, 2007). While the country experiences two wet and dry seasons, fewer than 15 percent of Kenya receives the rainfall needed to sustain crops (Library of Congress, 2007). This makes environmental issues, such as water shortages and diminishing water quality, a serious problem facing Kenyan farmers (Library of Congress, 2007).

In upcoming years, maize shortages in Africa are expected to total 10 million tons per year, a loss equal to five billion U.S. dollars (Edmeades, 2008). Of total maize loss, approximately 25 percent could be reduced with the creation of drought tolerant seed varieties (Edmeades, 2008). Public and private entities can help farmers manage their risk by creating drought tolerant seed varieties for specific location and climates in mind. Producers, on the other hand, can minimize their downside risk by adopting new seed varieties or implementing improved production inputs, in the form of fertilizers, pesticides, and irrigation equipment. However, all of these inputs can be costly, especially installation of irrigation systems, which include significant upfront cost that must be spread out over time. With sixty percent of Kenyans characterized as poor, chances to prosper agriculturally, without assistance, are limited (Library of Congress, 2007). However, with less than fifteen percent of the country receiving adequate rainfall, managing production risk through the use of drought tolerant seed varieties is an important component of production management that should be targeted by both private and public entities, as well as at the producer level (Library of Congress, 2007).

In recent years, several methods have been devised to improve maize production. The most significant component of these methods is the use of biotechnology to improve seed varieties to withstand biotic and abiotic stress conditions in eastern Africa. Biotechnology allows for higher yields as a result of improved resistance to droughts, pests, and weather stress (Nyabiage, 2009). This method provides the opportunity for farmers to reduce their risks and stabilize yields (Lybbert & Bell, 2010). A branch of biotechnology that has direct application to Kenyan farmers is the creation of drought tolerant seed varieties. While biotechnology crops protect against targeted risks, such as pests, they do not protect against drought. As a result, drought tolerant (DT) varieties are being created to provide a short term remedy for drought pressures in countries like Kenya. It can be argued that drought tolerant varieties are just as, if not more, important for Kenyans since drought affects all crops; whereas Bt varieties created to target certain pests may only effect one crop (Lybbert & Bell, 2010). While drought resistant varieties have a number of benefits, most importantly stable yields under moderate drought conditions, they may have costs in the form of reduced yields in extreme drought conditions in comparison to traditional varieties (Lybbert & Bell, 2010). In extreme drought conditions, DT varieties are undifferentiated from their non-DT varieties unless new varieties can also outperform existing ones (Lybbert & Bell, 2010).

In hopes of improving productivity and, ultimately, yields, the Kenya Agricultural Research Institute (KARI) has been researching bacillus thuringiensis (Bt) maize (Nyabiage, 2009). Despite the popularity of Bt crops and hybrid maize varieties, farmers in Kenya have been slow to adopt, especially with pesticide producers resistance to endorse these varieties (Nyabiage, 2009). There are a number of traditional varieties, which are often preferred by small farmers because of their familiarity with the seed. In 2006 and 2007, traditional -or open pollinated- varieties were estimated to be 1,700 tons in sales (Langyintuo et al., 2008). On the other hand, hybrid maize varieties accounted for 26,300 tons in sales in 2006 and 2007 (Langyintuo et al., 2008). Adoption rates of hybrid varieties in Kenya, for the same time frame, accounted for 74 percent of total maize area (Langyintuo et al., 2008). In comparison with surrounding countries, Kenya's adoption rate of hybrids is much higher than Tanzania and Mozambique, with only 22 percent of total maize area devoted to hybrid varieties each (Langyintuo et al., 2008). For those producers who have not adopted hybrid seed varieties, the major hindrance of adopting hybrid seed varieties is the high cost associated with obtaining the seed.

There are several key players, both private and public, participating in the effort to introduce and incorporate Bt and drought tolerant maize varieties into Kenyan farms' framework. The private sector is made up of seed companies which have the resources necessary to create seed varieties and then sell those varieties into suitable markets (Edmeades, 2008). A key player in the private sector, Monsanto, is the leader in Bt technology, producing Bt seed for vegetables, cotton, corn and oilseeds (Monsanto, 2010b). The company aims to release drought tolerant maize varieties in the U.S. market by 2012 (Edmeades, 2008). DT maize varieties would give farmers an eight to 22 percent increase in yields in drought conditions that typically reduce yields by as much as 50 percent (Edmeades, 2008). Other seed companies involved in the private sector include Pioneer Hi-Bred and Syngenta (Edmeades, 2008). However, private companies, like Monsanto, operate largely as an independent for-profit organization, earning revenue from seed licensing and trait technologies (Monsanto, 2010a).

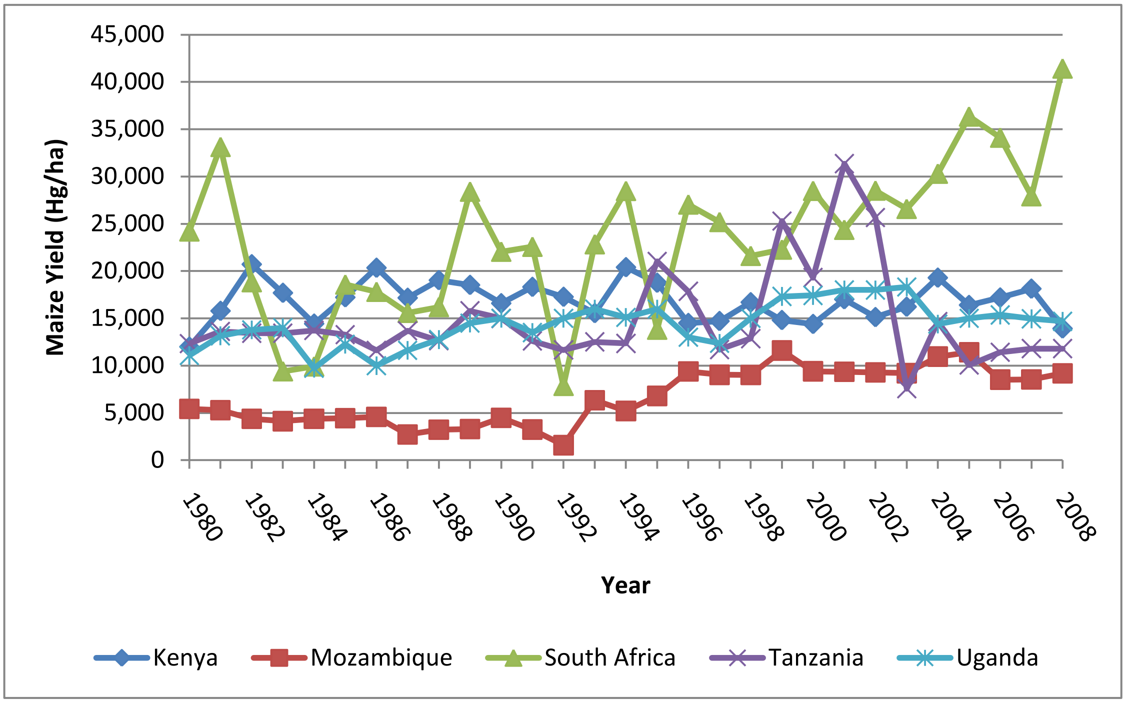
In order to get improved seed varieties from private seed companies to farmers who need them the most, the public sector acts as a facilitator (Edmeades, 2008). The International Maize and Wheat Improvement Center (CIMMYT) is a public entity that aims to improve the global maize and wheat innovation network to benefit the poor (CIMMYT, 2010). CIMMYT uses science to produce and employ productive varieties that can withstand adverse conditions, while at the same time providing a collection of resources, help farmers to implement conservation practices, and supplying the skills necessary to achieve all these objectives. Because of CIMMYT's efforts, with the aid of national programs, approximately 50 percent of hybrid varieties adopted are attributable to varieties developed through CIMMYT (CIMMYT, 2010). As a result of using hybrid varieties, 426 million hectares of crop area have been spared from drought loss (CIMMYT, 2010). ZM521, a drought tolerant seed variety released through a collaboration between private seed companies, CIMMYT, and the International Institute of Tropical Agriculture (IITA), is now planted on approximately one million hectares in Africa (Edmeades, 2008). As a result of CIMMYT's efforts, a project, Drought Tolerant Maize for Africa (DTMA), has been implemented to create improved drought tolerant varieties, aiding up to 40 million sub-Saharan Africans, and increasing the value of yields by up to $200 million through DT (Edmeades, 2008). The DTMA project was created by the public sector, specifically CIMMYT, and focuses on drought resistance efforts without using genetically modified seed. Specifically, this project targets both hybrid and open pollinated varieties. In addition, DTMA aims to target severe drought conditions, as well as low fertility characteristics.

Recent years have seen partnerships created among the public and private sectors to bring hybrid seed to those in developing countries who need it the most. For example, the Bill and Melinda Gates Foundation are collaborating with Monsanto to provide small farmers with access to drought tolerant seed varieties, improved tools and training, and access to markets (Gates Foundation, 2010). The Gates Foundation aims to improve productivity for small farmers' existing land and farm inputs, so that families can meet nutritional requirements while at the same time earning money from their cash crops (Gates Foundation, 2010). The result of this collaboration is a project, Water Efficient Maize for Africa (WEMA). This group makes use of Monsanto's technology expertise, CIMMYT's phenotyping and improved maize germplasm, and other entities, such as the African Agricultural Technology Foundation (AATF), which acts as a broker between the private and public sectors and farmers or agricultural institutions to insure that seed varieties are tested and then transported to various regions in Africa (Edmeades, 2008). By improving germplasms for specific areas in eastern and Southern Africa that are susceptible to droughts, the project aims to increase yields by 15 percent in drought conditions (Edmeades, 2008). Five countries (Kenya, Mozambique, South Africa, Tanzania and Uganda) are participating in this project. These drought tolerant hybrids should be available and usable beginning in 2017 and will all be available royalty free (Edmeades, 2008).

1.2 Motivation

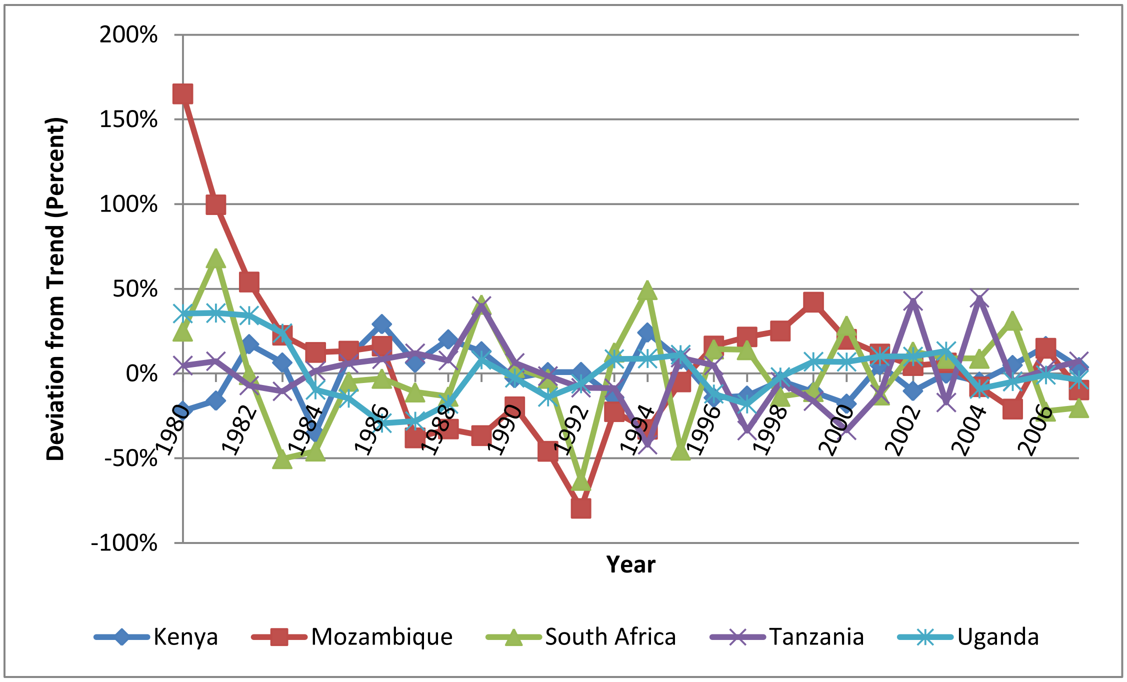
Although both projects, DTMA and WEMA target drought conditions, the projects do not assess variability in yields or the downside risk associated with production. Variability in yields can be a result of stochastic factors outside of the farmer's control, such as weather conditions, or a direct result of inputs chosen, such as seed variety. Despite the efforts made by the public and private sector to bring hybrid seed varieties to the poor, some farmers are still reluctant to adopt these varieties. Specific efforts should be made to educate farmers on the benefits and costs of hybrid seed adoption and how adoption can increase mean yields. While some farmers can learn simply by reading the additional benefits on the label, the majority require personal experience with a seed variety or recommendation from others' personal experience (Lybbert & Bell, 2010). It is also important that farmers understand that there are circumstances in which hybrid seed varieties might not operate optimally, such as, if inadequate fertilizer is used or in poor soil (Lybbert & Bell, 2010).

Figure 1-1. Maize Yields in Kenya, Mozambique, South Africa, Tanzania and Uganda



Source: FAO Stat, 2010.

Figure 1-2. Percent Deviation from Yield Trends in Kenya, Mozambique, South Africa, Tanzania and Uganda



Source: FAO Stat, 2010.

In eastern Africa, yields can vary across farms, and even within farmers' fields. As is shown in Figure 1-1, yields vary significantly from year to year within a country and have no general trend up or down over time. In Kenya, specifically, average yields ranged from a low of 12,000 hectograms per hectare to a high of 20,711 hectograms per hectare. Extreme highs and lows in total yields, within short spans of time are a pattern throughout the country, and show the instability within maize production. Figure 1-2 shows the percent deviation from the trend for each country, and displays proportionally low or high deviations away from the average yield trend for each country. Kenya and Uganda have the least deviation from the trend, while Mozambique and South Africa have high deviations from the trend. These figures show that variability exists in yields and can be attributed to a number of things, both within and beyond the farmer's control, such as access to inputs, infrastructures, and weather conditions.

This paper examines the variability associated with yields in Kenya, as well as the downside risk associated with production, specifically in drought prone areas. A stochastic production function is used as a model to incorporate factors of production in Kenya farms to analyze the magnitudes and causes of variability in yields among farmers. Yields will be analyzed for mean, variance and skewness. These models will be used to make conclusions about perceptions of traditional versus hybrid performance in both drought and non-drought prone areas of Kenya.

1.3 Thesis Organization

This thesis research is organized into chapters. Chapter 2 reviews literature pertaining to this thesis and is segmented into sections, which includes background information on yield response research, a summary of the hybrid versus traditional seed variety debate, and previous literature that concerns stochastic production function modeling. Chapter 3 discusses the data source. This includes information about the logistics of the data source, as well as descriptive statistics concerning the data broken down by areas of focus. It also describes the data in terms of inputs and outputs associated with the stochastic production function. Major categories of inputs, land, fertilizer, seed, and labor are discussed, as well as a host of other input characteristics. Output is also defined and described as it pertains to the data set. In addition, Chapter 3 includes descriptive statistics associated with the model. Chapter 4 presents the theoretical framework behind the model, including the functional form comparison and definition of all variables used for estimation. Chapter 5 presents the results of the ordinary least squares and weighted least squares regressions for the mean output response, the variance model, and the skewness model. It also analyzes how these models contribute toward the hybrid versus open pollinated/traditional varieties debate. Chapter 6 concludes this research by summarizing the model and data analysis results, while addressing the implications of the results and posing suggestions for future research.

Chapter 2 - Literature Review

For years, developing countries' have struggled to overcome weather, technological and resource barriers to provide sufficient food stocks. This struggle has been at the forefront of international agriculture. In sub-Saharan Africa, it is estimated that 300 million people identify maize as a staple in their diets (Edmeades, 2008). As such an important component of Kenyan's diets, farmers aim to produce enough maize to meet current demand; however, Kenya has been limited by maize shortages in recent years (Edmeades, 2008). These shortages are a direct result of stochastic factors, such as the dependence on rainfed agriculture and the region's susceptibility to drought risk. To combat these shortages, researchers have examined why such shortages exist in terms of efficiency among farmers, as well as stochastic factors associated with weather. Shortages are often a result of inefficiencies in production systems. These inefficiencies can exist for any number of reasons: such as, inefficient inputs, lack of access to capital or credit markets, or because of subsistence requirement demands. Subsistence requirements entail producing enough food to meet the needs of your own family's daily caloric intake.

Current studies make an effort to examine the reasons for variability in yields associated with agricultural productivity. Anderson and Hazell state several causes for yield variability, such as input variation, hybrid seed varieties, weather, technology and improved information (Anderson & Hazell, 1989). A trend in the research points to a farm's input mix as a source of difference in yields, often citing access to inputs as a determining factor of yield success (Edmeades, 2008). Additional research conducted by Nyoro, Kirimi, and Jayne aimed to evaluate the costs involved in maize production in both Kenya and Uganda and gave conclusions about production costs for farmers in Kenya (Nyoro et al., 2004). Overall, their study found that fertilizer, labor and land preparation costs make up the largest percentage of total production costs (Nyoro et al., 2004). As a result, cost differences are often a result of higher labor and fertilizer costs across different regions of the country. Cost differences also play a key role in the input mix, and, ultimately, in the explanation of yield variability.

2.1 Yield Response Data – Seed Varieties

There are a number of maize seed varieties available to Kenyan farmers. Often, differences in production costs are a significant result of seed type used in production, whether hybrid, traditional, recycled or new (Nyoro et al., 2004). Specifically, these seed varieties can be broken down into hybrid maize, open pollinated varieties, recycled open pollinated varieties, and recycled hybrid seeds (Nyoro et al., 2004). Hybrid maize seeds are varieties that have been bred to increase yields, specifically in areas that face issues such as drought, flooding, or pests. Recycled hybrid seeds are seeds from an existing hybrid plant that are replanted. These seeds often produce approximately 30 percent less yield than the original plant, but have the benefit of being inexpensive (Langyintuo et al., 2008). On the other hand, open pollinated varieties- better known as traditional – are seeds that are grown for their specific traits, that also have the ability to change, adapting to the environment (Primal Seeds, 2000). Unlike hybrid seeds, open pollinated varieties can be recycled or replanted while still maintaining seed traits. Thus, recycled open pollinated varieties. However, yields still decline approximately five percent with the recycling of open pollinate varieties (OPVs) (Langyintuo et al., 2008).

Farmers in Kenya continue to use recycled seed, despite extension agents' warnings of reduced yield, because of a belief that hybrid seed quality has decreased (Nyoro et al., 2004). Recycled hybrid seeds result in lower yield and often do not transfer on the traits that the hybrid seeds are used for, such as disease resistance or drought tolerance (Mabaya et al., 2009). Eventually, extension officials aim to expel the belief that hybrid seed quality has decreased and promote adoption of quality hybrids and certified seeds (Nyoro et al., 2004). Current reasons for lack of adoption include: limited access, high costs, purchase frequency, and lack of education about the benefits of hybrid seeds (Mabaya et al., 2009).

Looking at other inputs in Kenya, in terms of variability, fertilizer usage, irrigation, labor costs, and land preparation all contribute to differences in yields among farmers. Nyoro et al. (2004) show that fertilizer adoption rates in Kenya are steady. Another study finds that to achieve the highest yield potential, improved seed varieties must be paired with appropriate fertilizer type and usage (Jia, 2009). Although fertilizer usage has increased over the past ten years, there are still many factors that hinder fertilizer adoption rates. These reasons include: lack of credit, lack of access to fertilizer suppliers, inadequate irrigation, and little infrastructure. However, there are some factors working in favor of increased fertilizer usage which include marketing policies, shorter distances to retailers, and improved logistical infrastructures (Ariga, Jayne, Kibaara, & Nyoro, 2008).

High costs of production are associated with high input and machinery costs, poor extension systems, seed quality, and poor management decisions (Nyoro, 2004). Management decisions are a source of variability in production yields: as well, management decisions influence production. Any number of decisions made can affect yield quality and quantity. In the case of Kenya, management decisions often deal with water and soil technology. Seed technology adoption decisions often differ across regions, with large adoption rates in the high potential zones and low adoption rates in the arid zones (Bett, 2004). A key factor in management decisions lies in the understanding of the benefits associated with input adoption, both economically and socially (Bett, 2004). Farmers must understand that a proactive course of action will help them in the long run, even if there are upfront costs that burden a farmer's current financial situation. The concept of risk perception is intertwined with management decisions, specifically in Kenya, where inputs are major investments of farm financial capital. Individual risk is comprised of those risks that can be valued, such as water availability, and also any unforeseen events in the future combined with the individual's ability to handle the event (Doss, McPeak, & Barrett, 2005). Overall, Kenya is in a unique position with increasing hybrid maize adoption rates. With better seed quality and improved access to markets, Kenyan farmers will likely make the transition to hybrid seed varieties on their own, further increasing maize productivity throughout the country (Nyoro et al., 2004).

2.2 Traditional versus Hybrid Varieties

The traditional versus hybrid seed variety debate has been in existence for nearly a century, since the Pioneer Hi-Bred Corn Company introduced the first commercial hybrid seed onto the market (Pioneer Hi-Bred International, 2010). While hybrid seed varieties are typically recognized as the best quality of seed available, Kenyan farmers remain reluctant to adopt these varieties over open pollinated or traditional varieties. As areas across the world, such as Central America and Asia, adopt hybrid seed varieties into their production practices, Africa remains idle (Thurow, 2008). Reasons for traditional variety use over hybrids vary, but the most common cause is a belief that seed quality for hybrids has diminished in recent years (Nyoro et al., 2004). Kenyans may attribute poor production performance of hybrid varieties in drought prone areas, with diminished seed quality. Other reasons for using traditional varieties are that most open pollinated varieties have changed over time to adapt to their existing environment, as well as the ability to recycle seed (Nyoro et al., 2004).

However, on the opposite side of the spectrum, seed companies and extension specialists throughout the world are trying to encourage and implement the use of hybrid seeds to improve crop productivity. Proponents of hybrid seed varieties encourage adoption in Africa, especially as per capita food production decreases as a result of both stagnant yields and rapidly increasing population rates (Thurow, 2008). Scientists argue in favor of hybrid seed varieties and their ability to produce stable yields under stressed conditions. By using improved seed varieties, farmers decrease their downside risk, defined as the chance of loss associated with a particular outcome. Downside risk in agriculture is characterized by interactions between a number of nonlinear variables, for example, the interaction between yields and inputs, such as fertilizer or irrigation (OECD, 2009). In this case, fertilizer is a risk increasing input. Although risk increasing inputs simultaneously reduce risk associated with production yield under favorable conditions, under extreme unfavorable conditions, they can increase yield risk. On the other hand, irrigation is a risk decreasing input, meaning that the risk involved with not using the additional input is greater than the risk of using the input. In this case, risk decreasing inputs aim to increase overall output (Just, 2003).

Anderson and Hazell conclude that not only do hybrid varieties offer a greater chance of improved yields, but they are also less risky than traditional varieties (Anderson & Hazell, 1989). However, the debate for traditional varieties still holds in drought prone areas where the traditional varieties can perform better under extreme drought conditions than the hybrid varieties. In order to increase adoption of hybrid varieties in Kenya, seed providers, in collaboration with extension agencies, should work together to help provide information about the benefits of hybrids over traditional, as well as first hand experiences, which often prove to be more useful than other information provided to the farmer.

2.3 Stochastic Production Function

The stochastic production function was first introduced in 1978 by Just and Pope who made use of an input conditioned output distribution (Gardner & Rausser, 2001; Just, 1978). This function goes beyond the scope of classical inputs and allows for random elements associated with production uncertainty to enter the functional relationship. Further work contributed by

Aigner, Lovell and Schmidt used the stochastic production function framework as a process that includes a random element corresponding to inefficiencies in a firm's technical production as well as predicted elements (Battese, 1997). In this case, the function is no longer deterministic or explained within the model but also includes a variable to account for production uncertainty.

In production agriculture, a stochastic production function is used to account for random elements of production: such as, weather, price fluctuations and soil quality. The stochastic set-up also allows certain variables to be treated as deterministic, while incorporating random components (Gardner & Rausser, 2001). Random components, or stochastic factors, which are outside of the farmer's control, have been examined for decades and are of great interest for management and policy decisions (Gardner & Rausser, 2001). In the case of Kenya, a stochastic production function can be used to incorporate the underlying fact that production is often stochastic and depends upon outside random variables, such as rainfall. The use of a stochastic production function allows the random element to be analyzed despite any decisions made (Gardner & Rausser, 2001).

When looking at Kenya specifically, production uncertainty is a component of everyday farming practices. This uncertainty, especially when considering drought stress, is likely to impact the input choice mix (Babcock, 1995; M. Isik, 2003; M. Isik, 2002; Just, 1978; Pope, 1979; Ramaswami, 1992). Uncertainty must be taken into account when making decisions, both at the producer level, as well as at a policy level in distributing aid or designing policies to help combat production risk (Battese, 1997; Griffiths, 1982; Just, 1978; Pope, 1979). Uncertainty in Kenya is often a result of precipitation variation which can result in 20 percent lower or higher yields on an annual basis, making production risk for drought prone areas in Kenya a significant factor in the decision making process (Bullock, 1994; M. Isik, 2003) In 1997, Battese et al. argued that production risk was a major component lacking in the stochastic production function (Battese, 1997). Production risk can be made clear by analyzing the input mix and management decisions of a farm, and, therefore, should not be ommitted from the stochastic production function framework. Once inputs have been determined, the amount of specific inputs in the production function will result in a given output. These input differences can also account for differing yields from year to year (Kumbhakar, 2010).

Risk considerations or responses often explain many production decisions in a given timeframe. For example, diversification of crops will allow a farmer to continue farming despite various challenges, such as drought or pest problems (Di Falco, 2006; Naeem, 1994). In Kenya, where farmers are prone to drought conditions and lack irrigation capabilities, risk can be hedged by diversifying crops which will keep productivity steady and decrease downside risk exposure (Di Falco, 2006). Di Falco and Chavas concluded that, in particular, seed specific diversity not only helps depress environmental risk while sustaining productivity, but can also decrease variability of crop production for farmers applying small amounts of pesticide (Di Falco, 2006). While responses to risk, such as diversification, can be observed, it is often hard to identify if these responses are due to technology adoption, input limitations, financial constraints, or preferences towards averting risk (Just, 2003).

When examining stochastic production functions, it is also important to go beyond the first order mean moment, to higher moments of variance and skewness. First, examining previous research in the literature, Shankar et al. (2007) examined output risks for genetically modified crop technology, implementing a variance model to analyze the risk properties of Bt cotton. This study points out that while much research has been completed on the effect of Bt technology on increasing outputs or decreasing inputs, little knowledge is known about the effect on production risk associated with the technology. Shankar et al. (2007) implement two techniques to analyze the effect or preference of Bt technology over traditional: first-degree stochastic dominance analysis, as well as a stochastic production function regression incorporating a variance element. This study found that Bt technology increases yield risk, and increased yields, a result of genetic breeding to foster resilience, when growth conditions were favorable, but decreases yields when growth conditions were unfavorable (Shankar, 2007). Shankar et al. (2007) point out that while Bt technology is risk increasing, the improvement in mean yields justifies adoption or preference of this technology over traditional varieties.

Similarly, Traxler et al. (1995) analyzed the impact of varietal technology, specifically wheat technology, on mean and variance of yield. This study examined the impact of yield improvement at its height during the green revolution, and yield stability after the green revolution. Traxler et al. (1995) implemented a Just-Pope framework with a multiplicative heteroscedasticity model, where the variance of the error term can be linked to a number of independent variables. The authors aim to analyze varietal progress at differing nitrogen levels, since inputs, like irrigation and fertilizer, were factors in the success associated with yields in the green revolution (Traxler, 1995). A generalized least squares regression is used to examine mean and variance of yields. Traxler et al. (1995) find that in the years following the green revolution, there have been substantial gains in the release of varieties that increase yield, as well as promoting stability.

While Shankar et al. (2007) and Traxler et al. (1995) examine the variance of crop yields, Di Falco and Chavas (2006) focus their research on two higher order moments: variance and skewness. These authors aimed to examine the impact of crop genetic diversity on mean, variance and skewness of yield, as well as provide an analysis of the relationships between crop genetic diversity, risk management, and productivity. A stochastic production function is used to incorporate production risk into the model, following the approach by Antle (1983) and the Just-Pope (1978) framework, to incorporate elements of variance and skewness. The authors also analyze the effect of crop diversity upon welfare, by finding the impact of diversity on revenue, risk premium, and the certainty equivalent (Di Falco, 2006). In terms of results, the authors discover that crop genetic diversity can help increase farm productivity, manage risk, and reduce yield variability. Additionally, high crop genetic diversity levels have a positive impact on skewness of yields, indicating that diversity can reduce risk of crop failure (Di Falco, 2006). The techniques used by Di Falco and Chavas will be outlined and discussed further in Chapter 4 in relation to the variance and skewness models.

This research, similar to the previous literature, aims to determine whether yield variability exists and what sources of variability exist within maize production. In addition, this analysis will also make use of the framework used by Antle (2010), concerning potato production to evaluate partial order moments of skewness as a way to analyze asymmetric effects of inputs on yield distributions. The frameworks mentioned previously, will be used to analyze whether or not inputs are variance/skewness increasing or decreasing within the models.

2.4 Conclusion

Knowing that production shortages exist as a result of weather, technological, and resources barriers provides the opportunity for producers and consumers to find ways to prevent crop failure. Droughts are a factor outside of a producer's control; as a result, they experience variability from year to year unless preventative actions are taken to protect themselves from production risk. One of the best methods that farmers in Kenya can use to mitigate downside risk is adopting seed varieties that are designed to improve yields even under drought conditions.

However, Kenyan farmers have been slow to adopt these varieties for several reasons, such as expense, lack of familiarity with the product, and preexisting perceptions about hybrid seed quality. This research aims to look at the performance of seed varieties in Kenya. Yield data will be analyzed to determine the mean, variance and skewness among farmers in Kenya through the use of a stochastic production function model. Specifically, the analysis will look at the performance of traditional versus hybrid seed varieties in both drought and less drought prone regions of the country. This research also aims to determine whether the notions that hybrid seed varieties will decrease a farmer’s downside risk in drought conditions, and also if traditional varieties perform best under extreme drought conditions with little to no rain.

Chapter 3 - Overview of Data Set

This chapter describes the data set used for this research. Section 3.1 outlines the data source, identifying where the survey was conducted, who was surveyed, when the survey was administered, and what information was included in the data set. Section 3.2 outlines the data collection procedure looking specifically at where data was collected, and how the data was organized for analysis. Section 3.3 gives a summary of the data along with descriptive statistics for all crops surveyed within the data set. This information was used to narrow down the data into a specific sample for this research. Section 3.4 gives summary statistics about the following inputs: land, fertilizer, seed and labor, for all crops surveyed. These inputs were used to design a stochastic production function for data analysis. Section 3.5 lists and describes other input characteristics, beyond the classical scope, that were used in the stochastic production function framework. Section 3.6 lists inputs that were not included in the production function, as well as a justification for these not being included in the analysis. Section 3.7 describes the output associated with this research and how it was transformed to meet data analysis needs.

3.1 Data Source

The data used for this study was collected from a survey administered in Kenya in the following years: 1997, 2000, 2004, and 2007. The survey was designed and administered by Egerton University’s Tegemeo Institute of Agricultural Policy and Development and Michigan State University. The data was compiled into databases by officials at Michigan State University. The survey sampled 1,540 households across 24 districts. The resulting survey data consisted of information pertaining to household assets, crops produced, household demographics, fertilizer use, field level data, input use, labor specific inputs, and maize seed adoption rates and usage, with a total of 1,200 variables. The 2007 data set includes 1,397 households and is the focus of this research; however, it should be noted that descriptive statistics are provided for both the population sample of 1,397 households, as well as the narrowed sample size of 951 households which is used to model the mean, variance and skewness of yields approach.

The survey aimed to examine consumption and production patterns among rural households including, marketing outlets, input choice and adoption rates of seed varieties, specifically for maize. Objectives of the survey in relation to maize aspired to 1) discover causes for low technology adoption rates among farms, 2) determine what attributes and perceptions lead to production of differing maize varieties, 3) identify production practices, inputs, markets, and access to credit markets and extension services, and 4) identify strategies among households to minimize risk against outside threats (Muhammad, 2008). Information gathered from the data has been used to produce reports examining the following issues in Kenya: agricultural productivity, poverty reduction, access to markets, fertilizer trends, livelihood diversification, maize marketing, and the impacts of policy interventions (Ariga et al., 2008; Chamberlin & Jayne, 2009; Kimenju & Tschirley, 2008; Oehmke, Jayne, Aralas, & Mathenge, 2010).

3.2 Data Collection Procedures

To administer the survey, samples were chosen, based upon population, and then divided into agro-ecological zones, targeting the rural population to survey. Then, specific divisions were chosen according to population and cropping patterns. Next, locations and sub-locations within divisions were chosen by local officials. Household lists were then compiled, and households were chosen for interview according to a random pattern. Figure 3-1 shows geographically how many respondents came from each district within the sample of surveyed households specific to this research for maize plots only. Table 3-1 shows the number of household respondents by agro-regional and agro-ecological zones. Table 3-2 shows the breakdown of households surveyed by province and district. This gives a clear picture of how the data was narrowed down according to agro-ecological zones.

This research focuses on a subset of the whole data set, specifically, those zones that face more drought prone conditions, namely agro-ecological zones two (lowland), three (lower midland 3-6), and four (lower midland 1-2). Areas excluded from this analysis are those that experience high maize productivity. While the entire data set includes several crops, this analysis focuses solely on maize production. Further aggregations and synthesis of the data will be provided in later sections.

Table 3-1. Household Respondents by Zone for All Crops

Agro-regional zone

Name	Agro-ecological Zone	Households Surveyed	Percent
Coastal Lowlands	1	78	5.6
Eastern Lowlands	1, 3, 5	157	11.2
Western Lowlands	3	170	12.2
Western Transitional	4	157	11.2
High Potential Maize Zone	5, 7, 8	385	27.6
Western Highlands	6	147	10.5
Central Highlands	5, 6, 7	253	18.1
Marginal Rain Shadow	2	50	3.6
Total		**1397**	**100.0**

Agro-ecological zone

Name	Agro-ecological Zone	Households Surveyed	Percent
CL - Coastal lowland	1	88	6.3
L - Lowland	2	50	3.6
LM3-6 - Lower midland 3-6	3	279	20.0
LM1-2 - Lower midland 1-2	4	157	11.2
UM2-6 - Upper midland 2-6	5	271	19.4
UM0-1 - Upper midland 0-1	6	248	17.8
LH - Lower highland	7	263	18.8
UH - Upper highland	8	41	2.9
Total		**1397**	**100.0**

Table 3-1 shows that most respondents come from high potential maize zones and the Central Highlands. The table also shows the break down by agro-ecological zone. The largest percentages of respondents are from agro-ecological zones 3, 5, 6, and 7. This research focuses on zones 2, 3, and 4, eliminating high potential maize zones from the analysis and focusing on drought prone areas of the country. Table 3-2 shows the breakdown of households surveyed by province and district. This research focuses on households from the following districts: Bungoma, Kakamega, Siaya, Kisumu, Kitui, Laikipia, Machakos, Makueni, and Mwingi.

Table 3-2. Household Respondents by Province and District for All Crops

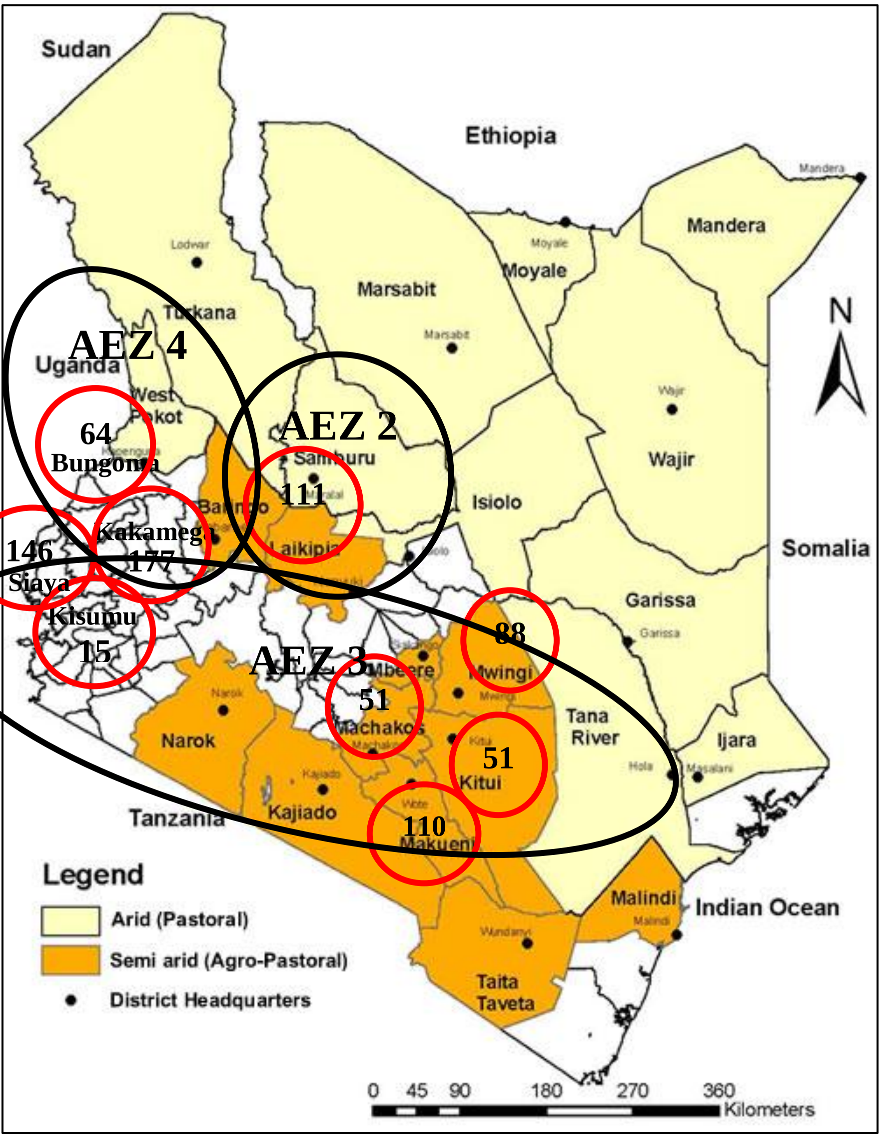
Province

Name	**Households Surveyed**	**Percent**	**Cumulative Percent**
Coast	88	6.3	6.3
Eastern	229	16.4	22.7
Nyanza	257	18.4	41.1
Western	279	20.0	61.1
Central	171	12.2	73.3
Rift Valley	373	26.7	100.0
Total	**1397**	**100.0**	

District

Name	**Households Surveyed**	**Percent**	**Cumulative Percent**
Kilifi	53	3.8	3.8
Kwale	25	1.8	5.6
Taita Taveta	10	.7	6.3
Kitui	18	1.3	7.6
Machakos	21	1.5	9.1
Makueni	75	5.4	14.5
Meru	82	5.9	20.3
Mwingi	33	2.4	22.7
Kisii	87	6.2	28.9
Kisumu	99	7.1	36.0
Siaya	71	5.1	41.1
Bungoma	82	5.9	47.0
Kakamega	137	9.8	56.8
Vihiga	60	4.3	61.1
Muranga	68	4.9	65.9
Nyeri	103	7.4	73.3
Bomet	39	2.8	76.1
Nakuru	103	7.4	83.5
Narok	24	1.7	85.2
Trans Nzoia	56	4.0	89.2
Uasin Gishu	101	7.2	96.4
Laikipia	50	3.6	100.0
Total	**1397**	**100.0**	

Figure 3-1. Respondents by Geographical Location for Maize Crops



Source: ASAL Based Livestock and Rural Livelihoods Support Project, 2008.

3.3 Data Summary and Statistics

The 2007 data used for this research is comprised of 1,397 households from 24 districts in Kenya. This data includes a number of variables pertaining to crop level production, input use, fertilizer use, labor, field level information, livestock production, and seed variety use. The frequency of each crop type is shown in Table 3-3. As is shown, maize is the predominant crop in Kenyan households with dry maize being produced on 11.2 percent of total fields, and green maize on 9.9 percent of total fields. The second largest crop share is attributed to beans which make up 9.9 percent of total field production. Intuitively, it makes sense that maize represents such a high share of production, as it is an important staple in the diets of Kenyan households. This research focuses solely on dry maize production in Kenyan agricultural households.

Table 3-3. Crops Produced

Crop	Frequency	Percent	Cumulative Percent
Maize-dry	2589	11.2	11.2
Beans	2284	9.9	21.1
Maize-green	2275	9.9	31.0
Sukuma wiki	1132	4.9	35.9
Bananas	970	4.2	40.1
Cowpeas leaves	915	4.0	44.1
Irish potatoes	766	3.3	47.4
Avocado	753	3.3	50.7
Napier / elephant grass	750	3.3	53.9
Indigenous vegetables / amaranthus	711	3.1	57.0
Sweet potatoes	680	2.9	59.9
Cowpeas	596	2.6	62.5
Mangoes	530	2.3	64.8
Onions	502	2.2	67.0
Guava	406	1.8	68.8
Cassava	396	1.7	70.5
Pumpkin	376	1.6	72.1
Pawpaws	343	1.5	73.6
Sorghum	320	1.4	75.0
Pumpkin leaves	308	1.3	76.3
Cabbage	302	1.3	77.6
Sugarcane-chewing	296	1.3	78.9
Lugard	293	1.3	80.2
Coffee cherries	243	1.1	81.2
Tomatoes	234	1.0	82.2
Passion fruit	234	1.0	83.3
Tea	226	1.0	84.2
Millet	213	.9	85.2
Indigenous grains	212	.9	86.1
Sugarcane	188	.8	86.9
Lemons	178	.8	87.7
Groundnuts	156	.7	88.3
Arrow roots	154	.7	89.0
Pigeon peas	153	.7	89.7
Oranges	152	.7	90.3

3.4 Land, Fertilizer, Seed and Labor Inputs

The inputs in the Kenyan data set are composed of the following major categories: land, fertilizer, seed type and labor. Land characteristics are composed of land size in acres, land tenure, land preparation, and land quality. Fertilizer characteristics are broken down into elemental components of nitrogen, phosphate, P_2O_5, and potassium from 26 different fertilizer formulations. Seed input characteristics are broken down into: 1) seed variety, specifically maize seed varieties, and 2) seed type, consisting of hybrid and OPV / traditional, and 3) quantity of seed used. Summary statistics for field level information of all crops are provided in Table 3-4.

Table 3-4 shows an interesting fact about household agricultural production in Kenya with 9,339 observations within the 2007 data. Since the total number of households for this data is 1,397, this shows that households have more than one field in most cases, with an average of 3.2 fields per household. In addition, the table shows that while field size varies, the largest percentage of fields are one acre or smaller, as most households own several small fields on which they farm. Next, looking at fertilizer use in 2007, the first thing that draws attention is that the total number of observations for fertilizer used does not match the number of fields total. In fact, this number shows that, approximately, only 58 percent of fields are using fertilizer. The next observation to point out in Table 3-4 is that nitrogen and phosphate usage rates are greater than potassium usage rates. For this reason, potassium use is dropped from the analysis, a topic that will be discussed later in this research. Last, the seed category in this table portrays an interesting picture of complex cropping systems in Kenya. The total number of observations for seed planted, 10,062 shows the presence of intercropping of different seed varieties on fields included in the data. In addition, Table 3-4 shows the importance of maize in Kenyan agricultural systems, with approximately 31 percent of total crops devoted to maize seed. This confirms the results of Table 3-3 and supports the focus of this research upon maize yield variability.

Table 3-4. Input Descriptive Statistics for All Crops

Input	Value	N	Mean	Standard Deviation	Minimum	Maximum
Land	Acreage under crops	9339	0.7	1.9	0.0	80
	Number of fields	9339	3.2	2.2	1.0	14
Fertilizer (kg/plot)	Fertilizer used (kg/plot)	5389	55.5	171.0	0.0	4500
	Nitrogen used (kg/plot)	5388	13.2	47.6	0.0	2070
	Phosphate used (kg/plot)	5388	12.3	47.4	0.0	1495
	Potassium used (kg/plot)	5388	1.2	12.2	0.0	460
Seed (kg/plot)	Seed planted	10062	123.3	1861.5	0.0	140000
	Maize seed planted	2879	9.3	21.5	0.1	700

Next, expanding further on fertilizer usage among maize plots, fertilizer was broken down into the elemental components of N (Nitrogen), P (Phosphate), and K (Potassium) for this research. Table 3-5 gives the descriptive statistics for each of these elements among maize plots. Examining potassium usage rates, on average zero kilograms are applied per acre with a maximum use of only 11.5 kilograms per plot of land. For this reason, potassium use was dropped from the list of independent variables in this research.

Table 3-5. Fertilizer Rate Descriptive Statistics for Maize Plots (kgs/acre)

Fertilizer Element	N	Mean	Median	Standard Deviation	Minimum	Maximum
Nitrogen Rate	951	4.1	0.0	9.8	0.0	124.0
Phosphate Rate	951	4.4	0.0	9.7	0.0	66.0
Potassium Rate	951	0.0	0.0	0.5	0.0	11.5

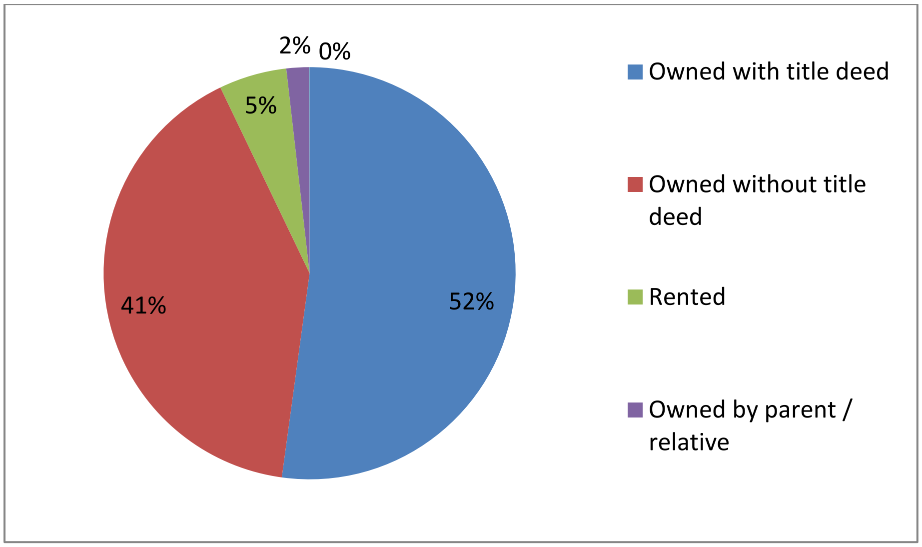
By contrast, Table 3-6 gives input descriptive statistics for land, fertilizer, and seed for maize plots analyzed in this research. First, the table shows that the average yield for maize plots is 459.1 kilograms per acre. Next, the average number of fields confirms the findings in Table 3-4, as households produce on an average of 1.5 fields. Fertilizer usage rates are broken down into nitrogen and phosphate, with potassium dropped from the data. Unlike the data for all crops, the number of maize fields and number of fertilizer observations is equivalent, showing fertilizer application for all fields. Last, maize seed planted is given in kilograms per acre. The table shows that, on average, 4.3 kilograms of maize seed are planted per acre.

Table 3-6. Input Descriptive Statistics for Maize Plots

Variable	Value	N	Mean	Median	Standard Deviation	Minimum	Maximum
Yield	Kilograms harvested per acre	951	459.1	360.0	425.6	0.0	3240.0
Land	Acreage under crops	951	1.2	1.0	1.2	0.0	10.0
	Number of fields	951	1.5	1.0	1.1	1.0	10.0
Fertilizer (kg/acre)	Nitrogen use	951	4.1	0.0	9.8	0.0	124.0
	Phosphate use	951	4.4	0.0	9.7	0.0	66.0
Seed	Seed planted (kg/acre)	951	4.3	4.0	2.3	0.0	16.0

In terms of land ownership, most fields are owned by the household either with or without the title deed. Figure 3-2 shows land percentage in terms of ownership, indicating whether the land is: 1) owned with title deed, 2) owned without title deed, 3) rented, 4) owned by parent/relative, or 5) government/communal/corporate owned. In terms of the first two options of land ownership, owned with/without title deed, this refers to whether or not the household has documented entitlement to their land, or whether they own the land without proper documentation. Rented land comprises a small amount (5.3 percent) of surveyed household’s ownership, while government, communal or cooperative lands are close to nonexistent among the households surveyed.

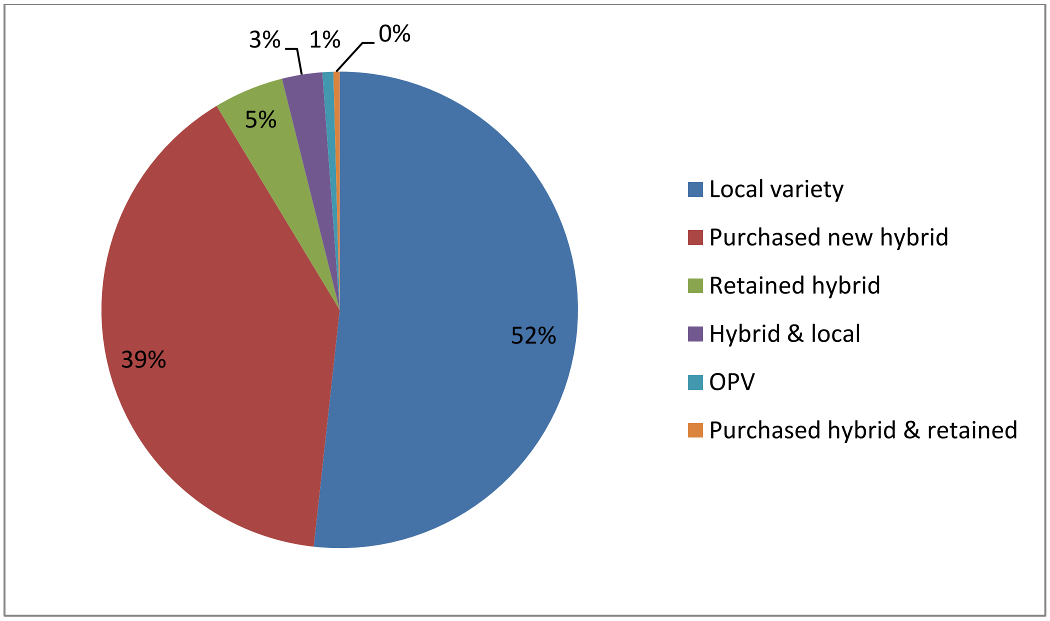
Figure 3-2. Land Ownership



N = 9,339

Seed quantity, type, and variety were all considered within this research. Seed varieties were narrowed down to maize seed, specifically. Seed types were broken down into local variety, purchased new hybrid, retained hybrid, hybrid and local, OPV, and purchased hybrid and retained. The break down by percentage for seed type is displayed in Figure 3-3, and shows that local varieties are the most prevalent, followed by purchased new hybrids.

Figure 3-3. Frequency of Seed Type for Sampled Households



N = 19,429

At this point, it is important to discuss aggregation of the independent variables within the data set. Figure 3-3 shows the importance of aggregating data into larger groups. For example, seed varieties are synthesized into two categories, hybrid and non-hybrids. The hybrids category consists of purchased new hybrid, hybrid and local, and purchased hybrid and retained categories into a single 'hybrid' category. On the other hand, OPVs and local varieties are grouped into a single 'non-hybrid' category. However, it should be noted that OPVs differ from local varieties with the difference between the two being that most local varieties are OPVs that have adapted to the local environment. In addition, labor and tenure were also aggregated for purposes of this data and will be discussed in further detail within this section.

Another key component of agricultural production in Kenya is labor inputs. In Kenya, labor consists of the following tasks: plowing, harrowing, planting, weeding, dressing, field dusting, stoking, harvesting, transport, shelling, post harvest dusting, bagging, storage, and security. However, the focus of this study is on labor activities up until the transportation of output, or the production activities undertaken by households for all pre-harvest activities and including harvest, these include: plowing, harrowing, planting, weeding, dressing, field dusting, stoking, and harvesting. Of these production activities, planting, weeding, and harvesting are completed most frequently among households.

Labor in Kenya can be carried out through hired labor, family labor or non-paid salary workers. The data used in this research breaks labor into: hired labor (days), adult females (hours), adult males (hours) and children (hours) providing labor. Table 3-7 gives an overview of how labor is broken down in terms of family and hired labor by activity for all households in the data set. The table shows the hours devoted to each labor activity by each household for the 2007 data. In comparison, Table 3-8 displays the breakdown of family labor, by male, female, and child labor, for all crops within the data set.

Table 3-7. Total Hired and Family Labor for All Crops

Family Labor (hours/activity/household)

Activity	Mean	Median	Standard Deviation	Minimum	Maximum
1st plowing	34.3	18.0	50.7	0.0	280.0
2nd plowing	29.7	20.0	31.6	0.0	120.0
Harrowing	13.0	6.0	17.2	0.0	70.0
Planting	29.0	20.0	35.0	0.0	396.0
1st weeding	50.0	30.0	69.1	0.0	528.0
Top-dressing	12.0	6.5	20.0	0.0	160.0
2nd weeding	42.3	28.0	54.0	0.0	448.0
Field dusting	15.7	8.0	26.8	0.0	180.0
Stoking	32.8	15.0	55.3	0.0	224.0
Harvesting	41.4	26.0	60.0	0.0	560.0

Hired Labor (days/activity/household)

Activity	Mean	Median	Standard Deviation	Minimum	Maximum
1st plowing	3.0	2.0	2.9	0.0	30.0
2nd plowing	5.0	4.0	4.5	0.0	14.0
Harrowing	4.0	1.5	5.4	0.0	12.0
Planting	1.6	1.0	1.1	0.0	8.0
1st weeding	2.9	2.0	1.9	0.0	14.0
Top-dressing	1.4	1.0	0.7	0.0	4.0
2nd weeding	2.9	2.0	2.0	0.0	18.0
Field dusting	1.6	1.0	0.9	0.0	4.0
Stoking	2.3	2.0	1.6	0.0	7.0
Harvesting	2.1	2.0	1.7	0.0	12.0

Table 3-8. Male, Female, and Child Family Labor for All Crops

Family Labor (hours/activity)

Male

Activity	Mean	Median	Standard Deviation	Minimum	Maximum
1st plowing	30.9	21.5	29.2	0.0	280.0
2nd plowing	25.7	20.0	21.6	0.0	120.0
Harrowing	13.5	6.5	16.5	0.0	70.0
Planting	19.0	14.0	18.1	0.0	160.0
1st weeding	36.9	25.0	37.3	0.0	288.0
Top-dressing	9.6	6.0	11.1	0.0	80.0
2nd weeding	30.8	24.0	28.0	0.0	224.0
Field dusting	11.1	7.0	12.6	0.0	60.0
Stoking	28.8	18.0	33.8	0.0	224.0
Harvesting	27.2	18.0	35.6	0.0	560.0

Female

Activity	Mean	Median	Standard Deviation	Minimum	Maximum
1st plowing	32.5	24.0	33.7	0.0	280.0
2nd plowing	22.8	18.0	19.4	0.0	120.0
Harrowing	14.5	15.5	9.0	0.0	30.0
Planting	17.8	12.0	22.1	0.0	396.0
1st weeding	37.9	24.0	44.1	0.0	528.0
Top-dressing	9.9	6.0	16.8	0.0	160.0
2nd weeding	31.8	21.0	36.5	0.0	448.0
Field dusting	19.4	10.0	30.2	0.0	180.0
Stoking	30.9	20.0	30.6	0.0	192.0
Harvesting	24.8	16.0	33.4	0.0	400.0

Children

Activity	Mean	Median	Standard Deviation	Minimum	Maximum
1st plowing	31.4	21.0	32.6	0.0	180.0
2nd plowing	18.3	12.0	17.6	0.0	72.0
Harrowing	6.3	5.0	3.2	0.0	10.0
Planting	16.9	12.0	16.3	0.0	96.0
1st weeding	35.0	21.0	39.6	0.0	252.0
Top-dressing	9.8	9.0	6.5	0.0	24.0
2nd weeding	29.8	20.0	31.1	0.0	224.0
Field dusting	12.0	11.0	5.9	0.0	20.0
Stoking	35.2	20.0	38.7	0.0	144.0
Harvesting	23.3	16.0	29.5	0.0	288.0

For simplicity, labor categories have been aggregated into: total hired labor (days) and total family labor (hours) for males, females, and children. Hired labor remains in terms of days worked and has not been converted into hours to mimic family labor hours since hired labor wages are computed with a daily rate. Table 3-9 shows the frequency of hired labor, male, female and child labor for all crops surveyed. This gives a clear picture of how labor is distributed among households.

Table 3-9. Frequency of Labor by Category for All Crops

Hired Labor

	Days per Household			
	0	**1 to 5**	**6 to 10**	**≥ 11**
Frequency	7815	1208	159	17
Percent	85.0	13.1	1.8	0.1

Male Labor

	Persons per Household			
	0	**1 to 5**	**6 to 10**	**≥ 11**
Frequency	3926	5129	130	14
Percent	42.7	55.8	1.4	0.1

Female Labor

	Persons per Household			
	0	**1 to 5**	**6 to 10**	**≥ 11**
Frequency	3145	5895	137	22
Percent	34.2	64.1	1.5	0.2

Child Labor

	Persons per Household			
	0	**1 to 5**	**6 to 10**	**≥ 11**
Frequency	8035	1140	23	1
Percent	87.3	12.4	0.2	0.1

For purposes of this research, the data is aggregated into two categories: total family labor and total hired labor. First, looking at hired labor, this type of input choice appears to be a rarity in agricultural production with 85 percent of fields per household reporting no hired labor. Looking closer at hired labor, data analysis finds a link between hired labor and the following activities: first plowing, planting, weeding, and harvesting. The data shows that these activities are often hired out by other parties in the event that the household does not have the means to complete the task. In terms of family labor, 57.3 percent of total fields had one or more male adults providing labor. Females also provide a source of family labor, and among the surveyed households, 65.8 percent of fields had at least one woman providing labor. As is common in agricultural production in developing countries, children also provide a form of family labor. However, in Kenya 87.3 percent of households reported no children providing family labor.

Within the sample of households analyzed for this research, labor was aggregated into hired labor (days per acre) and total family labor (hours per acre). Table 3-10 gives the descriptive labor statistics for households, and confirms the limited supply of hired labor discussed previously, with the average amount of hired labor being 2.8 days per acre. This also confirms the prevalence of family labor in Kenyan agricultural production with an average of 256.1 hours per acre, a minimum of zero hours per acre, and a maximum of 2,280 hours per acre.

Table 3-10. Descriptive Statistics for Labor Among Maize Plots

Labor Type	N	Mean	Median	Standard Deviation	Minimum	Maximum
Hired Labor (days per acre)	951	2.8	0.0	6.9	0.0	56.0
Family Labor (hours per acre)	951	256.1	182.0	285.7	0.0	2280.0

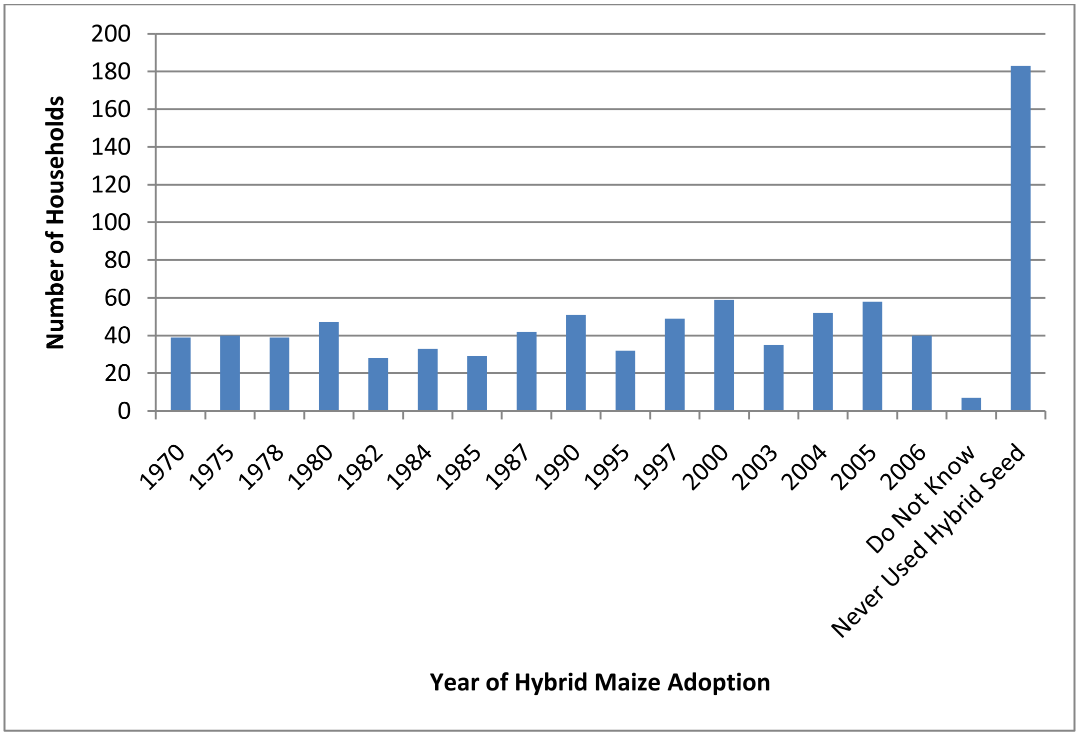
3.5 Other Input Characteristics

Other inputs that are considered within the data and this research include: harvest season, experience with hybrid maize seed, slope of the land, and access to credit. These inputs were chosen beyond the classical inputs to determine and further explain variability in maize yields within sampled households.

Figure 3-4 examines the observed frequency of experience with hybrid maize seed over approximately forty years. As the figure shows, only 0.5 percent of households indicated that they do not know about hybrid maize. This indicates that the knowledge about hybrid variety availability is present, but still approximately 14 percent of households have willingly chosen not to adopt these varieties. Despite lack of knowledge and unwillingness to adopt by households within the data, hybrid maize continues to be a trend whose adoption rates grow with experience. As Figure 3-4 shows, the highest record of observable hybrid maize adoption was in the year

2000, with 59 households adopting. In terms of cumulative percent, half of the sample who has adopted hybrid maize seed has done so since 1970, with the largest amount of observations in the 2000s. For many years, there has been debate concerning the effectiveness of hybrid varieties designed to combat specific known problems in regions compared to traditional varieties that have adapted to the environment on their own. This debate will be analyzed in the results and conclusions sections, in regard to the data analyzed to identify if any conclusions can be made given the results of this research.

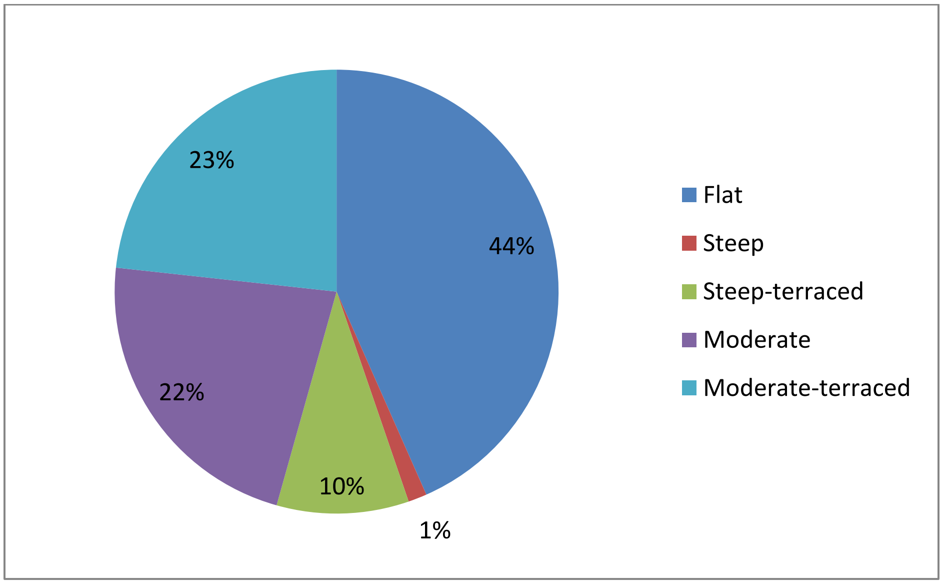
Figure 3-4. Experience with Hybrid Maize



N = 951

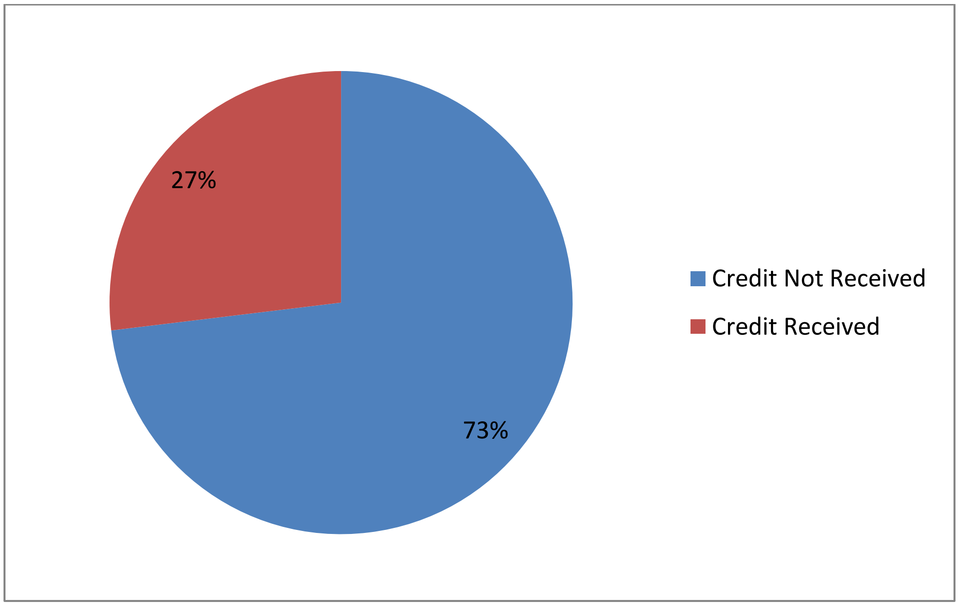
Next, slope of the land is also included in the model as an indicator of the quality of the land available to households. Figure 3-5 shows the percentage of observations for the following categories of land: flat, steep, steep-terraced, moderate, moderate-terraced. In terms of quality, these categories of land can be classified in terms of worst to best quality. For this research, specifically, terraced land is considered to be the most desirable for agricultural production among the sample. Terracing allows agricultural households to farm on sloping or steep land and also prevents soil erosion and runoff. In Kenya, terracing has been adopted as farmers see improved yields, a result of water remaining where it falls on terraced land (Critchley, 1991). As such, the presence of a variable explaining the effect that land slope has on maize yields further explains the variability among maize yields between households included within the data set.

Figure 3-5. Frequency of Slope Characterization, Percentage of Fields



N = 951

Figure 3-6. Frequency of Cash Credit Received for Maize Plots, Percentage of Households



N = 951

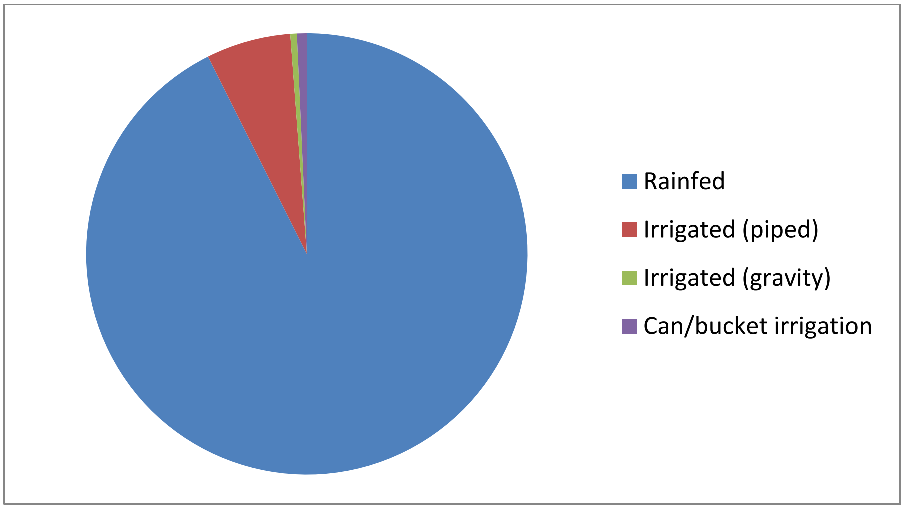
Access to credit was also analyzed in this research to determine whether or not this variable plays a role in yield variability among maize producers in Kenya. Figure 3-6 displays the breakdown of households within the sample who did or did not receive credit in the year 2007. As the figure shows, approximately 73 percent of households did not receive cash credit in 2007. This will be taken into consideration within the regression analysis to determine whether the influence of cash credit within the household had a positive or negative impact on maize yields for households.

Lastly, a final aspect analyzed in this research is the presence of multiple crops within households in the data set. This particular variable was included in the model to analyze whether there was variability in yields when looking at houses that produced only one crop compared to those that produced a portfolio of crops. However, in the case of the 2007 data set, all households produce more than one crop, indicating no variation among households for this variable. One explanation for this overwhelming amount of crop diversity among households is explained by the notion that households are largely subsistence. Households are, therefore, consuming at least a portion of what they produce and as a result must produce multiple crops to meet subsistence needs, and to also mitigate risk.

3.6 Physical Capital Inputs

Water source is also a key factor of crop production for Kenyan farmers; Figure 3-7 shows crop water systems available in Kenya according to their popularity. The largest percent of fields in Kenya, 92.6 percent, are rainfed. Irrigation (piped, gravity, and can or bucket) accounts for only 7.4 percent of all fields. Since the majority of fields, approximately 93 percent in the total data set and approximately 98 percent in the narrowed selection, consisted of rainfed agriculture it remains the focus of this research.

Figure 3-7. Crop Water Systems



N = 9339

Mechanization in this data set includes the ownership of assets which include: tractors, irrigation equipment, spray pump, car, water pump, ploughs for tractor, planter, harrow/tiller, truck, sheller, weeder, sprayer, and combine harvesters. Table 3-11 displays the frequency and percentage of households owning such assets. With only 9.4 percent of the households surveyed owning mechanization assets, it appears that this is an opportunity that has yet to be capitalized upon in Kenya. For the land area farmed, it may be too expensive to adopt mechanization assets, and expansion may not be feasible for a number of reasons. The households in the study may have lacked 1) access to mechanization opportunities, 2) funds to purchase and maintain mechanization (either cash or credit), and 3) knowledge or experience about how to use assets.

Table 3-11. Mechanization Assets

Asset	Frequency	Percent	Cumulative Percent
Spray pump	480	5.4	5.4
Irrigation equipment	137	1.6	7.0
Car	64	.7	7.7
Water pump	35	.4	8.1
Tractor	31	.4	8.5
Ploughs for tractor	28	.3	8.8
Planter	15	.2	8.9
Harrow /tiller	14	.2	9.1
Truck	9	.1	9.2

Sheller	8	.1	9.3
Ridger /weeder	5	.1	9.3
Boom sprayer	4	.0	9.4
Combine harvester	3	.0	9.4

Additional inputs that were included in the data set that have not been included for analysis are: fungicide, insecticide and herbicide. These items were not included since they comprised only 4.5, 7.7, and 2.5 percent, respectively, of households input use.

3.7 Outputs

Maize output in this research is the dependent variable, and is defined as yield, or kilograms harvested per acre. In the case of production, the amount harvested represents the output for a given amount of inputs, while yield, on the other hand, takes into account the production per unit of land for a given amount of inputs. For the purposes of this research, output is defined as yield to account for the varying sizes of land. This variable is used to examine yield variability among households that are a result of different input combinations, and to determine sources of variability.

3.8 Sample Model Descriptive Statistics

For purposes of this research, descriptive statistics for each agro-ecological zone were analyzed to identify the most drought prone areas within the selection of maize plots. In addition to identifying drought prone areas, the selection of maize plots were also divided into of hybrid seed varieties versus open pollinated and traditional varieties. The sample selection differs from the total and was narrowed down by crop produced, “maize-dry”, as well as scaled down to include only those districts that are considered to be the most drought prone within the data set, agro-ecological zones 2 through 4. These specifications brought the total from 1,397 households with 2,588 field observations to 459 households with 951 total field observations. From the set of 951 observations, the data was further broken down into hybrid, 406 observations, versus open pollinated and traditional varieties, 545 observations.

Table 3-12 examines the descriptive statistics for both output and inputs for all observations.

Table 3-12. Aggregate Descriptive Statistics for Maize Plots

Variable	Mean	Median	Standard Deviation	Min	Max	Variance	Skewness
Yield (kgs/acre)	459.1	360.0	425.6	0.0	3240.0	181166.9	2.0
Acres	1.2	1.0	1.2	0.0	10.0	1.4	2.7
Seed Quantity (kgs/acre)	4.3	4.0	2.3	0.0	16.0	5.4	1.3
Hired Labor (days/acre)	2.8	0.0	6.9	0.0	56.0	47.5	4.2
Family Labor (hours/acre)	256.1	182.0	285.7	0.0	2280.0	81634.5	2.8
Nitrogen Use (kgs/acre)	4.1	0.0	9.8	0.0	124.0	95.7	4.1
Phosphate Use (kgs/acre)	4.4	0.0	9.7	0.0	66.0	95.0	2.9
Harvest Dummy	0.6	1.0	0.5	0.0	1.0	0.2	-0.6
Hybrid Dummy	0.4	0.0	0.5	0.0	1.0	0.2	0.3
AEZ2 Dummy	0.1	0.0	0.3	0.0	1.0	0.1	2.4
AEZ4 Dummy	0.3	0.0	0.4	0.0	1.0	0.2	1.1
Terraced Dummy	0.4	0.0	0.5	0.0	1.0	0.2	0.4
Credit Dummy	0.3	0.0	0.4	0.0	1.0	0.2	1.0
Tenure Dummy	0.5	0.0	0.5	0.0	1.0	0.2	0.2
Years Hybrid Maize Experience	9.3	4.0	11.0	0.0	50.0	121.0	1.367

N=951, Standard Error of Skewness = .079

Comparing Agro-ecological Zones

Looking now at the synthesis of this data, information was analyzed to obtain the most drought prone areas within the sample to determine variability in maize yields. Those observations classified as high potential maize productivity zones were removed from the data set, as well as the following agro-ecological zones: 1 – coastal lowlands; 5 – upper midlands 2-6; 6 – upper midlands 0-1; 7 – lower highlands; and 8 – upper highlands. These areas were not included since this research focuses on drought prone areas. After synthesis, the following agro-ecological zones remain: 2 – lowlands; 3 – lower midlands 3-6; and 4 – lower midlands 1-2; remain with 111, 599, and 241 observations respectively. Table 3-13 gives an aggregated comparison of mean and standard deviation for all three agro-ecological zones.

Table 3-13. Agro-ecological Zones 2, 3, and 4 Descriptive Statistics for Maize Plots

Variable	AEZ 2 Mean	AEZ 2 Standard Deviation	AEZ 3 Mean	AEZ 3 Standard Deviation	AEZ 4 Mean	AEZ 4 Standard Deviation
Yield (kgs/acre)	362.9	317.4	342.4	323.8	793.4	509.2
Acres	0.9	0.5	1.4	1.4	1.0	0.9
Seed Quantity (kgs/acre)	5.0	2.6	4.2	2.4	4.3	2.0
Hired Labor (days/acre)	7.3	11.5	2.8	6.4	0.7	3.4
Family Labor (hours/acre)	326.6	369.4	194.4	191.6	377.0	377.6
Nitrogen Use (kgs/acre)	1.3	5.6	1.1	3.6	12.7	15.2
Phosphate Use (kgs/acre)	2.0	9.8	1.4	3.7	13.1	14.0
Harvest Dummy	0.6	0.5	0.6	0.5	0.7	0.4
Hybrid Dummy	0.5	0.5	0.3	0.5	0.7	0.4
AEZ2 Dummy	1.0	0.0	0.0	0.0	0.0	0.0
AEZ4 Dummy	0.0	0.0	0.0	0.0	1.0	0.0
Terraced Dummy	0.0	0.2	0.5	0.5	0.4	0.5
Credit Dummy	0.3	0.4	0.3	0.5	0.1	0.3
Tenure Dummy	0.7	0.5	0.4	0.5	0.4	0.5
Years Hybrid Maize Experience	13.5	10.7	5.6	7.7	16.6	13.5

Table 3-13 shows that in AEZ2 and AEZ3 that the average yield is 0.4 kilograms per acre for both districts. While AEZ4 has a higher average yield of 0.9 kilograms per acre, the zone still has areas highly susceptible to drought conditions with 90 percent of all observations producing approximately 1.5 kilograms per acre or less. The varying average production yield is likely a result of differing weather climates. Agro-ecological zone 2 is characterized by uneven rainfall distributions over the area, while agro-ecological zone 4 is classified as a wet area, overall, with varying temperatures within the zone. Agro-ecological zone 3, like agro-ecological zone 2 is characterized by inadequate rainfall over the area with both an increased intensity and frequency of drought.

As mentioned, the data for this research was further broken down into two categories: hybrid varieties and open pollinated and traditional varieties with 406 and 545 observations respectively. This break down allowed for inspection of whether hybrid varieties perform better than open pollinated or traditional varieties in drought prone areas, or whether there is truth in the argument that open pollinated traditional varieties have adapted to a region’s conditions to provide the most productive seed. Table 3-14 displays the mean and standard deviation descriptive statistics for all observations within the data, hybrid observations only, and non-hybrid only observations. A point to note, is the fact that hybrid varieties have over twice the average yield of non-hybrid varieties, and is an issue that will be brought up for discussion later in this chapter, as well as in the concluding chapter.

Table 3-14. Aggregate, Hybrid, and Non-hybrid Descriptive Statistics

	Aggregate		**Hybrid**		**Non-hybrid**	
Variable	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Yield (kgs/acre)	459.1	425.6	680.3	502.0	294.3	255.0
Acres	1.2	1.2	1.1	1.0	1.3	1.3
Seed Quantity (kgs/acre)	4.3	2.3	4.1	2.4	4.5	2.2
Hired Labor (days/acre)	2.8	6.9	3.1	7.4	2.6	6.5
Family Labor (hours/acre)	256.1	285.7	282.7	283.6	236.4	285.9
Nitrogen Use (kgs/acre)	4.1	9.8	8.3	13.5	0.9	2.7
Phosphate Use (kgs/acre)	4.4	9.7	8.7	12.9	1.3	4.2
Harvest Dummy	0.6	0.5	0.8	0.4	0.5	0.5
Hybrid Dummy	0.4	0.5	1.0	0.0	0.0	0.0
AEZ2 Dummy	0.1	0.3	0.1	0.4	0.1	0.3
AEZ4 Dummy	0.3	0.4	0.4	0.5	0.1	0.3
Terraced Dummy	0.4	0.5	0.4	0.5	0.4	0.5
Credit Dummy	0.3	0.4	0.2	0.4	0.3	0.5
Tenure Dummy	0.5	0.5	0.5	0.5	0.4	0.5
Years Hybrid Maize Experience	9.3	11.0	14.3	12.2	5.6	8.3

Chapter 4 - Methodology

This chapter discusses the theoretical model developed and describes the variables within the model. Previous literature shows that a stochastic production function is desirable to account for factors both within and outside of the household's control. Four production functions are developed within this model to determine amounts of variability among maize yields in Kenya, as well as look at the causes of variability among households. These production functions are: linear, quadratic, Cobb-Douglas, and two versions of the Generalized Leontief. Section 4.1 outlines the theoretical framework for the production functions used in this research. Section 4.2 looks at the different functional forms of production functions analyzed in this research. Section 4.3 will define and describe the variables within the model. Last, Section 4.4 outlines the model specification for the mean, variance and skewness regression models and discusses the results of each model.

4.1 Theoretical Framework

The idea that production functions may consist of not only classical inputs and outputs, but also can incorporate stochastic elements beyond the producer's control was developed by Just and Pope (1978). Further work was completed by DiFalco and Chavas (2006), whom provided an analysis of the associations between crop genetic diversity, risk management and farm productivity. DiFalco and Chavas also made an assessment of the impact of crop genetic diversity on the mean, variance and skewness of yield. Recently, Antle (2010) proposed methods to estimate asymmetric effects of inputs on potato yield distributions.

For this research, each production function framework is composed of both classical inputs and random components to account for production uncertainty, and one output, maize. In this research, output is measured in terms of yield, or tons harvested per acre. Classical inputs make up fourteen variables within the model. This can be shown by the following equation, specifying yield as a function of the following independent variables.

Yield = f (acres, seed quantity, hired labor, family labor, nitrogen, phosphate, season, seed type, agro-ecological zones, slope of land, credit assistance, land tenure, hybrid maize experience) (1)

Simplified, this model means that a producer faces a given production function represented by y = $f(\mathbf{x}, \mathbf{v}, \mathbf{e})$, where y is yield or output, **x** is a vector of inputs, and **v** is a vector of stochastic inputs, and e is a random production error. In this case, $f(\mathbf{x}, \mathbf{v}, \mathbf{e})$ represents the greatest possible output a producer can obtain with inputs **x** and **v**. In the analysis, **v** is considered to be a random vector, with a given probability distribution that is subject to change and represents production uncertainty or risk. This research, like DiFalco and Chavas (2006), looks specifically at the interactions between the inputs, **x**, and random variables, **v**. Shankar et al. (2007) build upon this production function framework by adding that **e** can be ordered from unfavorable to favorable conditions, where f_e $(\mathbf{x}, \mathbf{v}, \mathbf{e}) > 0$, where e is the partial derivative. The authors argue that risk averse producers will choose inputs, (**x, v**), in order to maximize $\int(f\ (\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{e}) - w_x x)$, where f represents maximum output given inputs **x**, **v**, and **e** and w_x represents input prices.

Building upon the Shankar et al. (2007) hypothesis, if **e** represents randomness of weather variability, then hypotheses can be made in regards to the risk increasing/decreasing nature of inputs, such as hybrid or OPVs/traditional varieties. For example, it is hypothesized that OPVs/traditional varieties are risk decreasing since they are considered to be environmentally stable and have adapted to their local environment. On the other hand, hybrids are hypothesized to be variance increasing, despite their promise of higher yields, as they create more variability in yields.

4.2 Functional Form Comparison

This research is broken into four production function models, as shown in Table 4-1: linear, quadratic, Cobb Douglas, Generalized Leontief (r=2), and Generalized Leontief (r=3). Just and Pope (1978) argue that common production function forms are restrictive in instances where risk plays a crucial role in production decisions. In addition, they argue that a production function should be flexible, so that input effects on the deterministic element of production are different than on the stochastic element. In 1979, Just and Pope add that traditional production functions are restrictive and imply that changes in inputs are directly related to changes in the variance of output, a restriction that contradicts other reports. To compensate for this contradiction a function must include two elements: one, determining the effect of input on expected output, and another explaining the effects of input on variability of output. This research aims to include both elements proposed by the Just and Pope framework.

Table 4-1. Functional Forms of Common Production Functions

Function	**Functional Form** (*i, j, k = 1, … , n*)
Linear	$y = \propto + \sum_i \beta_i X_i$
Quadratic	$y = \propto + \sum_i \beta_i X_i + \sum_i \sum_j \delta_{ij} X_i X_j$
Cobb-Douglas	$y = \propto \prod_i X_i^{\beta} i$
Generalized Leontief (r=2)	$y = \sum_i \sum_j \delta_{ij} X_i^{\frac{1}{2}} X_j^{\frac{1}{2}}$
Generalized Leontief (r=3)	$y = \sum_i \sum_j \delta_{ij} X_i^{\frac{1}{3}} X_j^{\frac{1}{3}}$

Source: Griffin et al. (1987)

Looking specifically at each functional form, it is evident that there are both strengths and weaknesses for each form. Each of these forms presents an algebraic formulation of a production function, with linear being the simplest, followed by the quadratic form which interacts each of the inputs with each other and relates maize yields to each of these inputs. According to Beattie et al. (2009), the complexity involved with a functional form depends upon the precision desired, as well as the production procedure. For example, the Cobb-Douglas form provides information about constant elasticity of substitution among inputs, while the linear form only allows for constant marginal productivities for each input. It is important to note that a translog function is missing from this analysis. While this functional form was initially included in the analysis, the presence of zero values interacting with one another limited its use.

4.3 Description of Variables

The production functions developed within this model consist of 14 independent variables that determine variability among the dependent variable, yield. Table 4-2 defines each variable in the data set and offers a description of how each was determined. The following variables were modified from the original data: total hired labor, total family labor, total nitrogen, and total phosphate. Both hired and family labor were aggregated across labor activities within households. Fertilizer was also broken down into elemental components of nitrogen and phosphate to provide a more accurate analysis of fertilizer application. As discussed earlier, potassium was dropped from the analysis as most households use compound fertilizers that have zero percent of this element.

Table 4-2. Description of Variables

Variable	Description
Acres	The size of land (field) in acres.
Seed Quantity	The quantity of seed used in tons per acre.
Hired Labor	The total amount of hired labor in days per acre. This variable is aggregated across production activities (e.g. first plowing, second plowing, harrowing, planting, first weeding, top-dressing, second weeding, field dusting, stoking and harvesting) within each household.
Family Labor	The total amount of family labor in hours per acre. This variable is an aggregation of total hours worked for men, women, and children across production activities (e.g. first plowing, second plowing, harrowing, planting, first weeding, top-dressing, second weeding, field dusting, stoking and harvesting) within each household.

Nitrogen Use	The total amount of nitrogen in tons applied per acre by each household. This variable takes fertilizers used by each household and breaks them down into the elemental form of nitrogen only.
Phosphate Use	The total amount of phosphate in tons applied per acre by each household. This variable also takes fertilizers used by each household (e.g. DAP) and breaks them down into the elemental form of phosphate only.
Harvest Dummy	Harvest is broken down into two seasons: main and short. The main season occurs from October 15 until January 20 in AEZ3, while the main season lasts from approximately March 15-20 until July 30 in zones AEZ3 and AEZ4, with exception to AEZ2 whose season only lasts until May 30. The short season occurs from March 15 to June 30 in AEZ3, while the short seasons last from October 15 until January 20 in AEZ3 and AEZ4. Again with exception to AEZ2, whose short seasons last until December 30. The harvest dummy takes a value of one for the main season and zero for the short season.
Hybrid Dummy	The type of maize seed used for each household consists of the following varieties: purchased new hybrid, retained hybrid, open pollinated varieties, local varieties, local seedlings / cuttings / splits, hybrid and local, purchased hybrid and retained, and IR maize. The hybrid dummy takes a value of one for purchased new hybrid, hybrid and local, and purchased hybrid and retained. The hybrid dummy takes a value of zero for retained hybrid, open pollinated varieties, and local varieties.
AEZ2 Dummy	The data consists of three agro-ecological zones in Kenya: AEZ2 or the lowlands, AEZ3 or the lower midland 3-6, and AEZ4 or the lower midland 1-2. The AEZ2 dummy takes on a value of one for

	those households in AEZ2, and a value of zero for those households in AEZ3 and AEZ4.
AEZ4 Dummy	The data consists of three agro-ecological zones in Kenya: AEZ2 or the lowlands, AEZ3 or the lower midland 3-6, and AEZ4 or the lower midland 1-2. The AEZ4 dummy takes on a value of one for those households in AEZ4, and a value of zero for those households in AEZ2 and AEZ3.
Terraced Land Dummy	This data consists of five different slope of land characterizations: flat, steep, steep-terraced, moderate, and moderate-terraced. The terraced land dummy takes on a value of one for land characterized as steep-terraced and moderate-terraced, and a value of zero for land characterized as flat, steep, or moderate.
Credit Dummy	The credit dummy takes on a value of one if the household received cash credit that they tried to obtain, and a value of zero if they did not receive the cash credit that they tried to obtain.
Tenure Dummy	This data consists of five types of land ownership or tenure: owned with title deed, owned without title deed, rented, owned by parent / relative, government / communal / cooperative. The tenure dummy takes on a value of one if the land is owned with or without a title deed or owned by parent/relative, and takes on a value of zero if the land is rented or owned by government / communal / cooperative.
Hybrid Maize Experience	The number of years the household has had experience with hybrid maize as of 2007.

4.4 Model Specification

Shankar et al. (2007) proposed a production function regression examining further moments of yield, such as variance, as a function of input. This research aims to go beyond the scope of typical production function regressions where yield is a function of a set of inputs, by examining further moments, specifically, variance, and skewness similar to the approach used by Shankar et al. (2007), Di Falco and Chavas (2006), Traxler et al. (1995), and Antle (2010). Agricultural has long been susceptible to production risk, and inputs impact output variability (Shankar, 2007).

First, the Just-Pope has been modeled by many authors over the years and the stochastic production function is given by

$$y = f(\mathrm{H}, \mathbf{X}, \alpha) + g\,(\mathrm{H}, \mathbf{X}, \beta) + h\,(\mathrm{H}, \mathbf{X}, \delta\,) + \varepsilon_i \tag{2}$$

Where, H represents a set of dummy variables, (1 = purchased new hybrid, hybrid and local, and purchased hybrid and retained. Seed, 0 = retained hybrid, open pollinated varieties, local varieties, local seedlings / cuttings /splits, and IR maize). **X** is a vector of production inputs, and α, β, and δ are vectors of parameters that will be estimated, while ε is a randomly distributed error term. The first term in equation 2 represents the mean model approach of the stochastic production function, while the second and third terms represent the variance and skewness models, respectively. The first term does not include risk implications beyond marginal productivities.

Following the framework proposed by Traxler et al. (1995), this specification uses an exponential form, exp (H, **X**, β)) for g (H, **X**, β), and exp (H, **X**, δ)) for h (H, **X**, δ). In this case the absolute value of the residual is squared to find variance, where g = exp $|e_i|^2$, and the absolute value of the residual is cubed to find skewness where h = exp $|e_i|^3$. Next, estimation was achieved using a weighted least squares (WLS) mean regression of $y = f(\mathrm{H}, \frac{1}{\widehat{\sigma_i}}\,\mathbf{X}, \alpha) + \varepsilon_i$. For the variance and skewness models, OLS regression was used. This provides updated estimates that are both asymptotically efficient and consistent.

For this research, an ordinary least squares (OLS) regression will be completed and analyzed. If the OLS regression results are homoskedastic, these results will be discussed;

however, if the results are found to be heteroscedastic then a weighted least squares (WLS) regression will be completed and discussed in place of the OLS mean model. Next, variance and skewness OLS models will be created and used to determine sources of variability within the model.

The OLS regression was carried out on the data set of household field level information for the 2007 production data. The production function regression was performed on yield with acres, seed quantity, hired labor, family labor, nitrogen use, phosphate use, and years of hybrid maize experience included as explanatory variables. Dummy variables were included for harvest season, hybrid variety, agro-ecological zones, terraced land, credit, and tenure. Interaction terms were included between all inputs and between the hybrid dummy and all inputs. The majority of interaction parameters in the mean models were significant. An F-test found that the interaction terms were jointly significant. The variance and skewness models do not include the interaction terms for simplicity.

Models of Full Order Moments

This research aims to go beyond the traditional scope of classical production function models by including models of full order moments, specifically, variance and skewness. Shankar et al. (2007) focus their research on examining output risks for genetically modified crop technology, implementing a variance model to analyze the risk properties of Bt cotton. Similiarly, Traxler et al. (1995) analyzed the impact of genetic improvement on mean and variance of yield. On the other hand, Di Falco and Chavas (2006) focus their research on two full order moments: variance and skewness. Di Falco and Chavas aimed to examine the impact of crop genetic diversity on mean, variance and skewness of yield. This research, similar to Di Falco and Chavas (2006), looks at all three moments: mean, variance, and skewness of yield, but also makes use of the empirical modeling followed by Shankar et al. (2007), Traxler et al. (1995), and Antle (2010).

Variance Model

The following regression models, following DiFalco and Chavas, as well as Traxler et al. (1995) were considered to examine variance and skewness among yields. From the initial OLS regression or mean yield approach, the residual, u_i is retained. It is then manipulated to create the dependent variable for both the variance and skewness models. In this case, the ordinary least squares residual is used to estimate the marginal effects of explanatory variables on the variance and skewness of maize production.

For this research, two moments of the distribution were taken, and this section outlines the theoretical framework of the variance model. DiFalco and Chavas, who conducted a similar study examining the links between crop genetic diversity, farm productivity, and risk management, argue that the idea of risk aversion means that a mean-preserving increase in the variance of yield makes farmers worse off (Di Falco, 2006). In this research, variance relates specifically to how far maize yields are spread out from each other for individual household fields, and gives a clear picture of whether inputs are variance increasing or decreasing. This analysis examines how far the household field level observations lie from the mean. Building upon the mean model, the variance equation model, specifically, is given below in equation 3:

$$\ln |u_i|^2 = g\ (\mathrm{H}, \mathbf{X}, \beta) + \varepsilon_\mathrm{i} \tag{3}$$

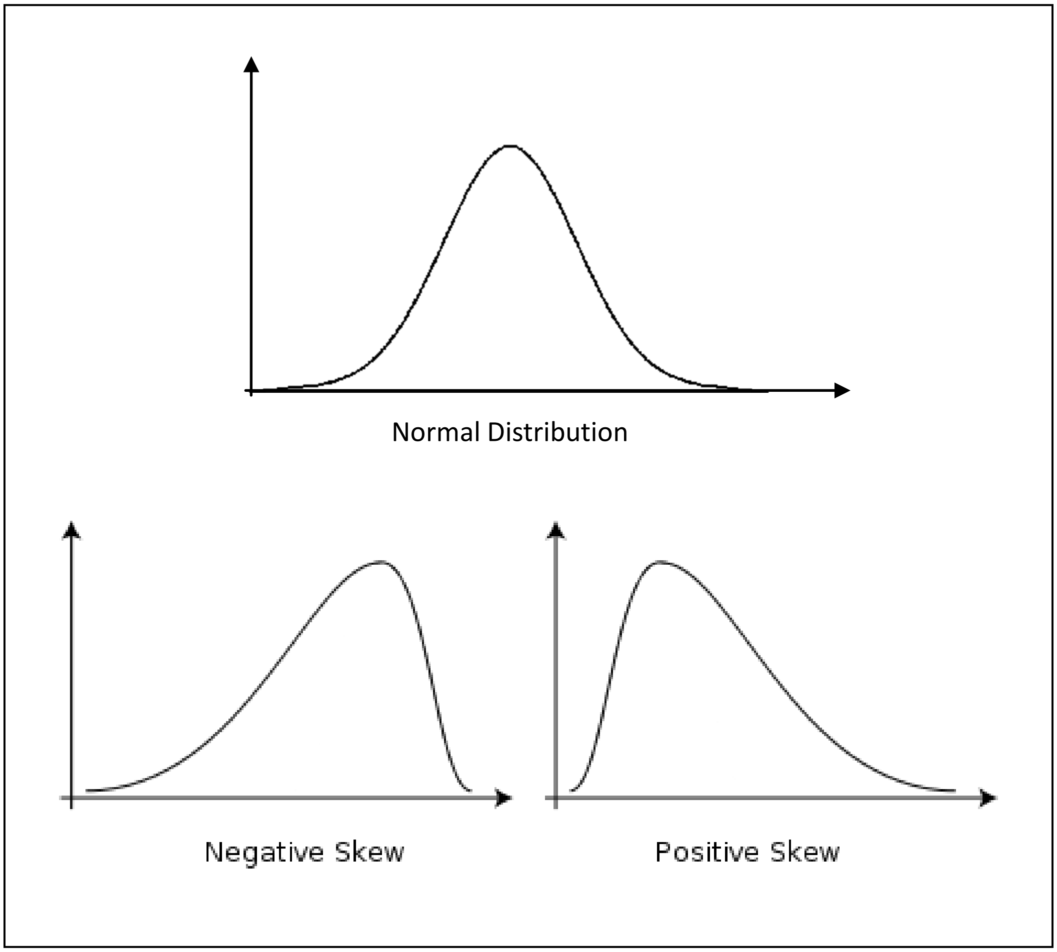
In terms of risk aversion, this research aims to build upon the idea that a risk averse farmer is made worse off by mean-preserving increases in the variance of yield, as well as the idea that farmer wellbeing is affected by the skewness of yield (Di Falco, 2006). DiFalco and Chavas argue that the majority of farmers display decreasing absolute risk aversion (DARA) (Binswanger , 1981; Di Falco, 2006). Applying this idea directly to this research, this would indicate that risk averse farmers have an incentive to plant seed varieties that reduce the variance of yields and positively affect skewness of yields, ultimately limiting their downside risk exposure. Downside risk exposure in this research relates specifically to the prevalence of drought susceptibility in Kenya leading to crop failure (Di Falco, 2006).

Skewness Model

Skewness in statistical theory looks at the asymmetry of probability distributions for random variables. This measure can be negative, where the tail on the left side of the distribution is longer; positive, the tail on the right side of the distribution is longer; or zero, where tails are either symmetrical or nearly symmetrical. Skewness of yield in this research relates to the level of exposure to downside risk.

Figure 4-1 displays the normal probability distribution, as well as examples of positive and negative skewness. In the normal probability distribution, all observations are centered around a mean. In a negatively skewed distribution, the majority of the distribution is concentrated on the right side of the distribution with the left tail being longer. In a positively skewed distribution, the majority of the distribution is concentrated on the left side of the distribution with the right tail being longer.

Figure 4-1. Asymmetry of Probability Distributions



Downside risk is the probability that crops will fail when conditions worsen, such as worsening drought conditions resulting in failed crops. This measure, like variance, also indicates whether inputs are skewness increasing or skewness decreasing. Hardaker et al. (2004) point out that downside risk can also occur when a risky outcome depends upon non-linear interactions between a number of random variables. From this model, the research aims to indicate whether farmers adopt inputs that positively affect skewness, and therefore, decrease downside risk exposure. The skewness model resembles the variance model and is listed below in equation 4. To reiterate, u_i is the residual derived from the initial OLS regression on mean yield.

$$\ln |u_i|^3 = h\,(\mathrm{H}, \mathbf{X}, \delta\,) + \varepsilon_\mathrm{i} \tag{4}$$

Further expanding on the concept of skewness, zero skewness presents a normal distribution, centered around a mean. A positive or negative impact on skewness implies that the distribution is not normal and that the impact works to shift the distribution. In equation 4, the skewness model gives no indicator of positive or negative skewness impacts as the absolute value of the residual is taken. This approach, taken by Di Falco and Chavas (2006), does not account for positive and negative impacts, and signs for coefficients in the variance and skewness models are the same with different magnitudes. Antle (2010) goes on to explain the importance of analyzing partial order moments of skewness to identify asymmetric effects of inputs on yield distribution. For this analysis, both the full order moment of skewness, as well as the positive and negative partial order moments of skewness will be discussed.

Partial Order Moments

Antle (2010) justifies partial moment functions as a way to examine asymmetric effects of a number of inputs on yield distribution. Antle points out that key information can be lost by focusing only on full moments, such as variance and skewness. For this reason, this research builds upon the Antle skewness model by examining partial order moments, for both positive and negative skewness. This section identifies how asymmetric distributions were created and displays how the partial moments of skewness signify yield distributions and their relationship to input choice. The partial order moment of skewness for positive and negative skewness are represented by equations 5 and 6, respectively.

$$\ln u_i^3 = h\,(\mathrm{H}, \mathbf{X}, \delta\,) + \varepsilon_\mathrm{i} \quad \text{for } \varepsilon_\mathrm{i} > 0 \tag{5}$$

$$\ln u_i^3 = h\,(\mathrm{H}, \mathbf{X}, \delta\,) + \varepsilon_\mathrm{i} \quad \text{for } \varepsilon_\mathrm{i} < 0 \tag{6}$$

To summarize, first, the mean model was estimated using an OLS regression. From this model, the residual, u_i, is retained and both squared and cubed to create the dependent variable for subsequent models. In addition to the full order moment skewness model, the partial order moment models are estimated to identify factors affecting positive and negative deviations within the sample. This allows for determination of whether input effects are asymmetric in the skewness model. Last, a WLS regression was carried out for the mean model approach to correct for heteroskedasticity.

Chapter 5 - Results

This chapter presents the results from the econometric approaches discussed in Chapter 4. First, a cumulative distribution analysis between hybrid and non-hybrid seed technology is presented in Section 5.1 to identify general tendencies about the impact of the hybrid seed technology. This research uses production function regression on mean, variance and skewness of output on inputs to further analyze the effect of hybrid varieties on yield variability. The mean OLS and WLS regression results are presented in Section 5.2, including a discussion of implications of the findings. Section 5.3 discusses marginal productivity figures associated with the mean model. Section 5.4 presents the results of the variance model results, and includes a discussion of the impact of different inputs on variance. Next, Section 5.5 presents the findings from the full order moment skewness model, while Section 5.6 presents the findings from the partial order moments of skewness, examining positive and negative deviations. Section 5.7 discusses the hybrid versus OPV/traditional varieties in relation to the mean, variance and skewness models.

5.1 Cumulative Distribution Analysis

According to Shankar et al. (2007), stochastic dominance techniques provide a way to analyze risk associated outcomes when preferences are unknown. First-degree stochastic dominance (FDSD) is a technique that assumes only that producers prefer more to less. In this research, this implies that producers prefer higher yields to lower yields, a reasonable assumption. Comparing two technology prospects graphically, if the cumulative density function (CDF) of one lies to the right of another for any given probability level, then the one lying to the right is preferred to the one on the left. If either of the CDFs cross, they cannot be evaluated using FDSD, and second-degree stochastic dominance (SDSD) must be evaluated. SDSD goes beyond the scope of FDSD in assuming that 1) more is preferred to less and 2) producers are risk averse (Shankar, 2007).

Figure 5-1 presents a graphical interpretation of a cross tabulation for seed varieties by agro-ecological zone. In two of the three districts the frequency of hybrid varieties is greater than OPV/traditional varieties. However, in AEZ3, OPV/traditional varieties far outweigh hybrids.

Figure 5-1. Cross Tabulation of Seed Variety by Agro-ecological Zone for Maize Plots

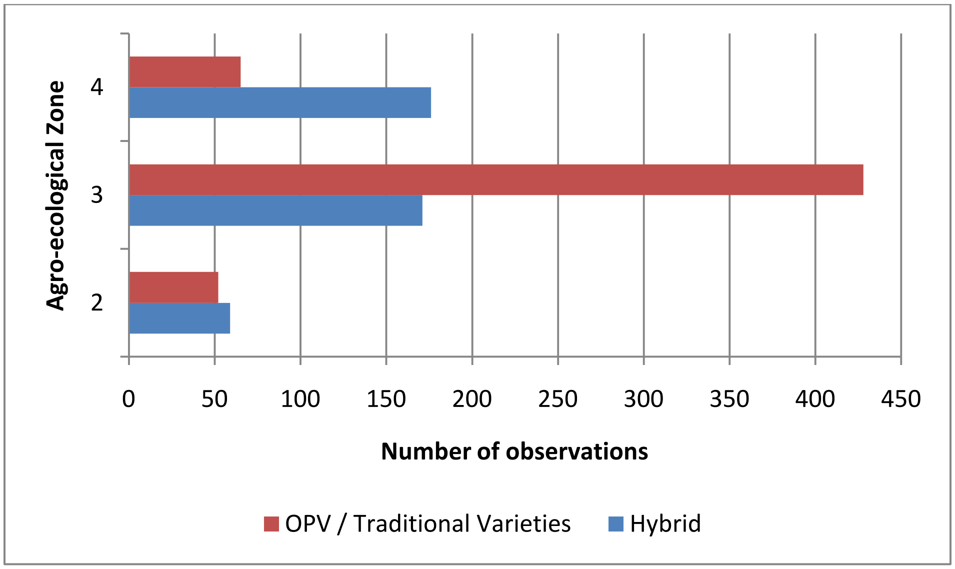
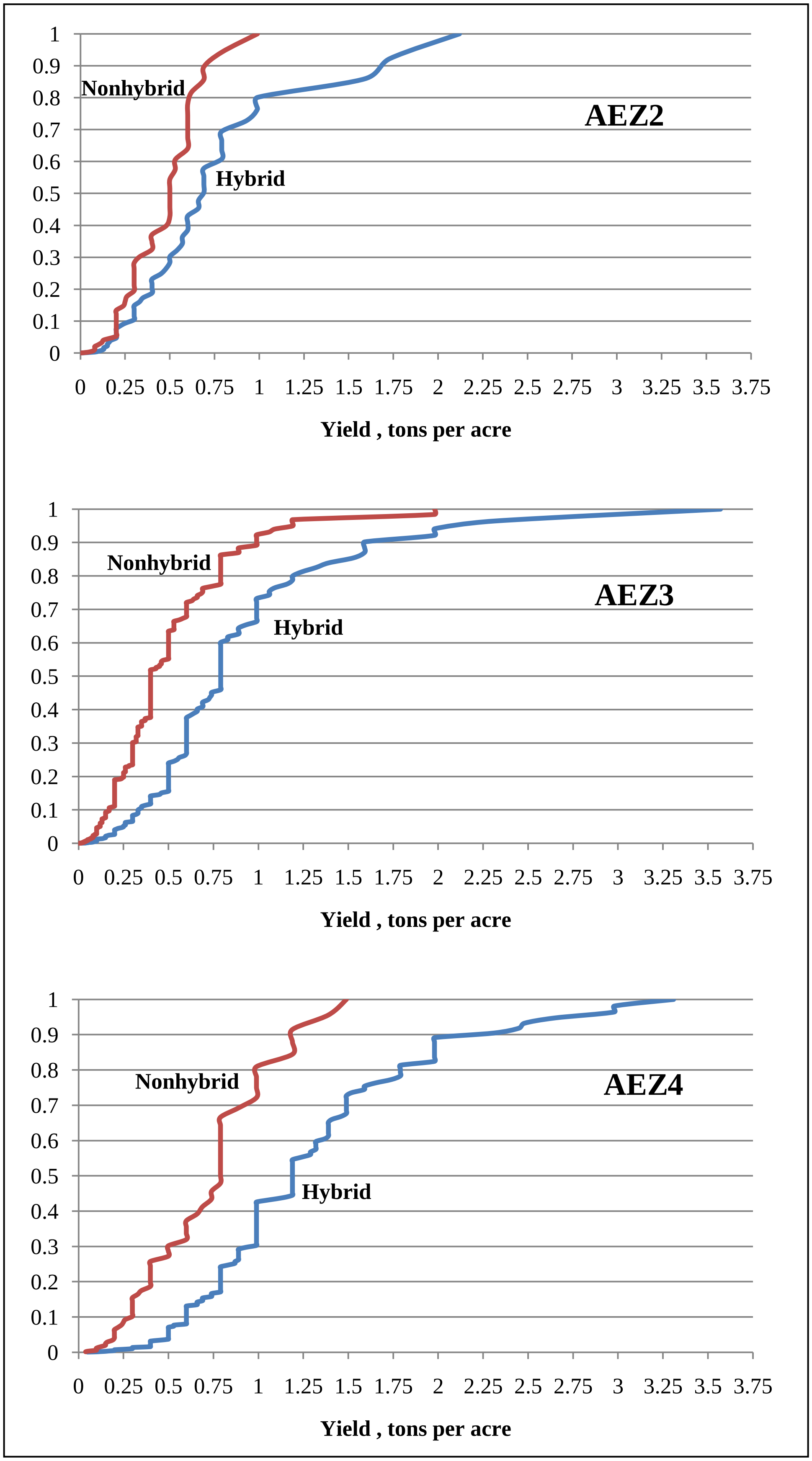


Figure 5-2 displays the cumulative distribution of yields for hybrids versus non-hybrids in each AEZ. Aggregation of all zones shows that the yield associated with hybrid varieties is higher than yield associated with non-hybrid varieties in all cases. According to Shankar et al. (2007), this fact should bring awareness to the idea that hybrid varieties, while having the ability to reduce the probability of crop failure, also may not reduce the probability of failure enough to decrease the probability of low returns. The risk associated with low returns, is a key factor in countries, like Kenya, who are averse to downside risk (Shankar, 2007).

An interesting point of Figure 5-2 is discovered when examining low yield outcomes for both hybrid and non-hybrid seed technologies. In this case, it is important to account for the cost of the hybrid seed technology although the figure does not reflect net returns associated with the technologies. The average seeding rate is 10 kilograms per acre in Kenya. Kenyans pay, on average, 400 Kenyan shillings (Ksh) for two kilograms of seed. If a producer aims to plant 10 kilograms of seed, this results in an additional seed cost of 2,000 (Ksh) for 5 bags. For producers to be willing to adopt hybrid seed, they would need to see a 65-90 kilogram increase in maize yields. In Kenya there is a need for policy making hybrid seed varieties more easily available to producers who face income

Figure 5-2. Observed Cumulative Frequency of Yields for AEZ 2, 3, and 4



constraints. Despite the increased yields associated with hybrid varieties, some producers still continue to not adopt. This may explained by the cost of hybrids in comparison to OPVs and traditional varieties, indicating that a price subsidy may be needed to provide incentive to farmers to switch.

While a stochastic dominance analysis is useful, limitations exist, primarily the assumption that results diverge only because of technological differences or randomness (Shankar, 2007). However, the data makes clear that input choices among households and even at the field level differ, creating divergences in the yield distribution pertaining to neither technological differences nor randomness (Shankar, 2007). In the case of stochastic dominance, without added price information, this technique will fail to tell the complete story and cannot fully discriminate on the preference of one variety over another. In addition, stochastic dominance analysis indicates the presence or use of both hybrid and non-hybrid seed varieties among one household observation; however, this is not the case with the data set. For these reasons, this research does not pursue a complete analysis of stochastic dominance. However, in order to better identify the effects of hybrid varieties on yield and variability, an alternate approach must be designed to control for differences. Similar to Shankar et al. (2007), this research uses production function regressions of mean and variance of output on variable inputs to further analyze the effect of hybrid varieties on yield variability. Next, building upon Di Falco and Chavas (2006), Traxler et al. (1995), Antle (2010), this research also uses production function regressions of skewness of output on variable inputs to further analyze sources of variability.

5.2 OLS and WLS Mean Output Response

The first step OLS estimates are included in Table 5-1 for all functional forms analyzed in this research while Table 5-2 presents the OLS regression results for the interaction effects in the quadratic and Generalized Leontief forms. While both tables are a product of one single regression, the results were split into two tables out of necessity. These OLS estimates are included, despite inefficiencies caused by heteroscedasticity, to display the differences between the OLS and WLS regressions.

After looking at the mean regression results, a Breusch-Pagan test was conducted. The test statistic is distributed as chi-squared under the null hypothesis of homoscedasticity. The values of these tests were: 314.0 for the linear form ($P \leq 0.001$), 293.6 for the quadratic ($P \leq 0.001$), 223.8 for the Cobb-Douglas ($P \leq 0.001$), 322.8 for the generalized Leontief (r=2) ($P \leq$

0.001), and 337.3 for the generalized Leontief (r=3) ($P \leq 0.000$). As a result of heteroscedasticity, the least squares estimators are biased and inefficient; the estimates of the variances are biased, as well, which creates problems and invalidates tests of significance. It should be noted that the White robust standard errors have been calculated and are included in Table 5-1 to present robust OLS estimates for comparison. In this case the presence of a robust standard error indicates that it is reliable when the regression errors are heteroscedastic and also results in more coefficients being significant.

Table 5-1. OLS Regression Results for All Functional Forms for Maize Plots

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**
INTERCEPT	0.1	***	4.4E-02	-0.1		0.1	0.8	*	0.4	-0.5	**	0.2	-1.2	**	0.5
Acres	-2.9E-02	***	9.9E-03	4.3E-02		3.8E-02	-0.3	***	0.1	0.3	***	0.1	0.8	***	0.3
Seed Quantity (tons/acre)	21.4	***	5.9	72.0	***	22.1	0.1	**	0.1	11.8	***	3.5	9.6	***	3.2
Hired Labor (days/ acre)	3.4E-04		2.0E-03	-9.3E-04		7.8E-03	-0.1		0.1	0.0		0.1	-3.6E-02		0.1
Family Labor (hours/acre)	1.8E-05		6.4E-05	2.2E-04		2.0E-04	-0.1	***	3.1E-02	0.0	*	0.0	0.1		0.1
Nitrogen Use (tons/acre)	8.3	**	4.1	11.3		12.4	0.2	***	4.8E-02	-2.3		3.7	-1.4		2.9
Phosphate Use (tons/acre)	7.3	**	3.6	-4.0		11.4	3.6E-02		4.6E-02	1.3		3.7	-4.4E-02		2.8
Harvest Dummy	0.2	***	2.1E-02	0.2	***	2.1E-02	0.7	***	0.1	0.2	***	2.1E-02	0.2	***	2.1E-02
Hybrid Dummy	0.2	***	3.0E-02	0.3	***	0.1	0.5	***	0.1	0.4	*	0.2	0.4		0.3
AEZ2 Dummy	-0.1	***	3.9E-02	-0.1	***	4.1E-02	-0.5	***	0.2	-0.1	**	4.0E-02	-0.1	**	3.9E-02
AEZ4 Dummy	0.1	***	4.5E-02	0.1	**	4.5E-02	0.3	***	0.1	0.1	***	4.4E-02	0.1	***	4.4E-02
Terraced Dummy	-0.1	***	2.8E-02	-0.1	***	2.9E-02	-0.3	***	0.1	-0.1	**	3.0E-02	-0.1	**	3.1E-02
Credit Dummy	-1.4E-03		2.5E-02	1.7E-02		2.5E-02	-0.2	**	0.1	0.0		2.4E-02	1.3E-02		2.4E-02
Tenure Dummy	0.1	**	2.4E-02	0.1	***	2.3E-02	0.1		0.1	0.1	***	2.3E-02	0.1	***	2.3E-02
Years Hybrid Maize Experience	2.4E-03	*	1.4E-03	2.6E-03	*	1.4E-03	8.7E-03	**	3.7E-03	2.8E-03	**	1.4E-03	2.8E-03	**	1.4E-03
Model Performance															
N	951			951			951			951			951		
R^2	0.45			0.51			0.35			0.50			0.49		
Adjusted R^2	0.44			0.48			0.34			0.48			0.47		
F Statistic	39.35	***		21.18	***		41.8	***		20.61	***		20.56	***	
Heteroskedasticity Test (chi^2)	313.98			293.59			223.79			322.80			337.26		

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

Table 5-2. OLS Regression Results for Interaction Effects for Maize Plots

	Quadratic			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**
INTERCEPT	-0.1		0.1	-0.5	**	0.2	-1.2	**	0.5
Acres*Seed Quantity	-4.0		11.4	0.0		0.0E+00	0.0		0.0
Acres*Hired Labor	0.2	***	2.1E-02	0.3	***	0.1	0.8	***	0.3
Acres*Family Labor	0.3	***	0.1	11.8	***	3.5	9.6	***	3.2
Acres*Nitrogen Use	-0.1	***	4.1E-02	3.2E-03		0.1	-3.6E-02		0.1
Acres*Phosphate Use	0.1	**	4.5E-02	1.6E-02	*	8.5E-03	0.1		0.1
Seed Quantity*Hired Labor	-0.1	***	2.9E-02	-2.3		3.7	-1.4		2.9
Seed Quantity*Family Labor	1.7E-02		2.5E-02	1.3		3.7	-4.4E-02		2.8
Seed Quantity*Nitrogen Use	0.1	***	2.3E-02	0.2	***	2.1E-02	0.2	***	2.1E-02
Seed Quantity*Phosphate Use	2.6E-03	*	1.4E-03	0.4	*	0.2	0.4		0.3
Hired Labor*Family Labor	2.1E-03		3.5E-03	-0.1	**	4.0E-02	-0.1	**	3.9E-02
Hired Labor*Nitrogen Use	-1625.2		1397.8	0.1	***	4.4E-02	0.1	***	4.4E-02
Hired Labor*Phosphate Use	4.6E-05		1.5E-04	-0.1	**	3.0E-02	-0.1	**	3.1E-02
Family Labor*Nitrogen Use	5.9E-08		8.7E-08	1.3E-02		2.4E-02	1.3E-02		2.4E-02
Family Labor*Phosphate Use	-129.5	*	67.5	0.1	***	2.3E-02	0.1	***	2.3E-02
Nitrogen Use*Phosphate Use	366.5	*	191.4	2.8E-03	**	1.4E-03	2.8E-03	**	1.4E-03
HybridDum*Acres	-17.1	***	4.6	-5.9	***	1.6	-5.3	***	1.6
HybridDum*Seed Quantity	-1.1E-03		4.7E-03	-5.8E-05		2.7E-02	2.8E-02		0.1
HybridDum*Hired Labor	-2.2E-04	**	9.8E-05	-9.7E-03	**	4.0E-03	-3.6E-02	*	2.1E-02
HybridDum*Family Labor	-3.2		4.0	-0.8		2.0	-1.0		1.5
HybridDum*Nitrogen Use	6.0	*	3.4	1.7		1.7	1.8		1.3
HybridDum*Phosphate Use	-0.5		0.7	-0.4		0.5	-0.3		0.5
Acres^2	0.0		0.0						
Seed Quantity^2	4.3E-02		3.8E-02						
Hired Labor^2	72.0	***	22.1						
Family Labor^2	-9.3E-04		7.8E-03						
Nitrogen Use^2	2.2E-04		2.0E-04						
Phosphate Use^2	11.3		12.4						

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

Since the standard deviation of the error term is not constant over all values of the explanatory variables, the WLS regression helps correct for lack of equality in the error variances. In this case, different observations are not treated equally, and the WLS regression gives different weights proportional to $1/\widehat{\sigma_i}$ to different observations. The following equation is used to derive the WLS regression results for this model:

$$y_i = f_1(\mathrm{H}, \frac{1}{\widehat{\sigma_\mathrm{i}}} \boldsymbol{X}, \alpha) + \varepsilon_\mathrm{i} \qquad (7)$$

Equation 7 takes the original OLS regression and uses the new proportional data to estimate the WLS regression. This regression is presented to show the effect that weighting a key variable, such as land size, has on the results. While the regression coefficients do not change much, the standard errors are different. The WLS regression corrects for heteroscedasticity and also leads to more efficient unbiased estimates. The regression estimates from the WLS regression are presented in Table 5-3 and Table 5-4, and was split into two tables out of necessity.

Table 5-3 shows the results in which the model is weighted by the first variable, acres, giving those observations with less land size greater weight. The regression results show that at least 9 or more of 14 variables, were statistically significant at the one percent level. Looking at the model, the WLS regression results are both similar and different from the OLS mean model results in many ways. First, after correcting for heteroskedasticity, the R^2 for each model increased from a range of 0.35 to 0.51 in the OLS model, to a range of 0.34 to 0.63 in the WLS model, with most forms displaying a higher R^2 than its OLS counterpart. In addition, correcting for heteroskedasticity also increases the number of significant variables within each functional form.

In terms of similarities, this model also examines the inverse field size productivity hypothesis, indicating that additional acreage has a negative impact on maize yields, similar to the findings in the OLS mean model regression. Next, looking at the impact of harvest season, the results show that this variable is significant at the one percent level across all functional forms. Looking at AEZ4, consistent with the OLS results this variable is also positive and significant across all forms. Observations within this zone experience higher mean yields as they are located in less arid districts, where rainfall is more frequent. Last, as with the results of the OLS regression, years of hybrid maize experience does not have an impact on average maize yield. The dummy for hybrid varieties is positive in all models and significant across four models at the one percent level.

In contrast, the OLS model differed in a number of cases from the WLS model. For example, hired labor, which was insignificant in all forms in the OLS model, is only significant at the one percent level in the linear and Cobb-Douglas functional forms. Hired labor, specifically, has mostly a negative impact on yields. Now, looking at family labor, the results in the OLS model only found this variable to be significant in the Cobb-Douglas and Generalized Leontief (r=2) form; however, the WLS model shows that while this variable is significant at the one percent level across all functional forms, family labor has a small impact on average yield. This indicates that there is a slight overuse of labor within the Cobb-Douglas model. In the linear and Cobb-Douglas forms, family labor is negative, which may indicate that family labor is being used optimally at the field level. Next, AEZ2 is positive and significant in the linear and generalized Leontief forms. This contradicts the results of the OLS regression where AEZ2 was negative. Now, examining the terraced land dummy, the results show that this variable is significant in most forms, and positively impacts yields. This result differs from the OLS results presented in Table 5-1 where terracing has a negative impact. Last, the effect of land tenure was analyzed, for those whose land was either owned by the household or a relative, the results were significant in three forms, and had a positive impact on maize yields.

Now, examining the effects of fertilizer use on yields the regression results support the hypothesis that additional nitrogen use increases yields, but does not support the hypothesis that additional phosphate use increases yields. The interpretation of the nitrogen use coefficient in the linear form is that, ceteris paribus, one additional ton per acre of nitrogen increases mean yield by 19 tons per acre. This fertilizer usage rate can be used to calculate the expected price of fertilizer, by taking the equation, $P \cdot MP = r$, where P is maize price, MP is marginal product, and r is fertilizer price rate; in this case, 19 is substituted for MP, showing that producers should use more fertilizer if the price is less than 19 times the price of output. Fertilizer is often an input that is beneficial in favorable weather conditions and unfavorable in severe drought conditions, where a lack of water prevents plants from getting the nutrients they need to survive.

Table 5-3. Weighted Least Squares Regression Results for All Functional Forms

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**	**Coeff**		**Std. Error**
INTERCEPT	2.9E-02		4.9E-02	-0.2		0.1	0.6		0.5	-1.1	***	0.3	-2.4	***	0.6
Acres	-0.2	***	4.9E-02	0.3	**	0.2	-0.3	***	0.1	0.9	***	0.3	1.9	***	0.5
Seed Quantity (tons/acre)	49.6	***	5.0	49.5	*	29.8	0.3	***	0.1	14.6	***	3.4	12.9	***	3.1
Hired Labor (days/acre)	-6.4E-03	***	1.4E-03	-2.4E-03		7.3E-03	-0.1	***	3.4E-02	-4.4E-02		0.1	-0.1		0.2
Family Labor (hours/acre)	-1.0E-04	***	3.2E-05	4.1E-04	***	1.5E-04	-0.1	***	3.2E-02	2.5E-02	***	8.3E-03	0.2	***	0.1
Nitrogen Use (tons/acre)	19.0	***	2.8	51.6	***	20.4	0.4	***	0.1	8.3		6.0	7.8		5.1
Phosphate Use (tons/acre)	-8.6	***	2.4	-36.0	***	13.7	-0.3	***	0.1	-9.2	*	5.0	-8.8	**	4.5
Harvest Dummy	0.2	***	3.0E-02	0.2	***	2.8E-02	0.6	***	0.1	0.2	***	2.8E-02	0.2	***	2.8E-02
Hybrid Dummy	0.3	***	3.1E-02	0.3	***	0.1	0.5	***	0.1	0.5	**	0.2	0.4		0.3
AEZ2 Dummy	0.4	***	4.5E-02	0.1		4.8E-02	1.8E-03		0.1	0.1	*	4.9E-02	0.1	**	4.9E-02
AEZ4 Dummy	0.2	***	3.6E-02	0.1	***	3.5E-02	0.2	*	0.1	0.1	***	3.5E-02	0.1	***	3.5E-02
Terraced Dummy	0.1	***	3.5E-02	0.1	*	3.3E-02	0.1		0.1	0.1	***	3.4E-02	0.1	***	3.3E-02
Credit Dummy	-0.1		3.8E-02	-1.0E-02		3.6E-02	-0.2	**	0.1	8.1E-03		3.6E-02	1.4E-02		3.6E-02
Tenure Dummy	0.1	**	3.0E-02	3.6E-02		2.8E-02	0.1		0.1	0.1	**	2.8E-02	0.1	*	2.8E-02
Years Hybrid Maize Experience	8.5E-04		1.5E-03	9.2E-04		1.4E-03	1.2E-02	***	3.5E-03	1.4E-03		1.4E-03	7.3E-04		1.4E-03
Model Performance															
N	951			951			951			951			951		
R^2	0.51			0.60			0.34			0.63			0.63		
F Statistic	70.75	***		33.45	***		34.1	***		44.19	***		44.47	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

Table 5-4. Weighted Least Squares Regression Results for Interaction Effects

	Quadratic			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error
INTERCEPT	-0.2		0.1	-1.1	***	0.3	-2.4	***	0.6
Acres*Seed Quantity	-49.2	***	18.0	-12.0	***	3.4	-9.6	***	2.7
Acres*Hired Labor	2.7E-03		9.7E-03	2.3E-02		0.1	-1.3E-02		0.1
Acres*Family Labor	3.1E-06		2.3E-04	2.8E-03		8.5E-03	-1.5E-02		3.8E-02
Acres*Nitrogen Use	-12.9		10.7	-8.9	**	4.0	-7.5	**	3.3
Acres*Phosphate Use	14.7		9.7	9.1	***	3.6	7.9	***	3.0
Seed Quantity*Hired Labor	-2.7E-02		0.4	-2.4E-04		0.4	0.1		0.5
Seed Quantity*Family Labor	-3.9E-02	***	1.1E-02	-0.3	***	0.1	-0.7	***	0.2
Seed Quantity*Nitrogen Use	-3782.7	***	1523.6	-33.3		42.8	-16.0		16.4
Seed Quantity*Phosphate Use	1151.5		1119.5	5.6		36.2	5.7		14.3
Hired Labor*Family Labor	7.4E-06	***	2.6E-06	1.8E-03	***	7.1E-04	9.7E-03	*	5.1E-03
Hired Labor*Nitrogen Use	-0.1		0.8	2.3E-02		0.6	-0.2		0.6
Hired Labor*Phosphate Use	0.3		0.8	0.3		0.6	0.5		0.5
Family Labor*Nitrogen Use	-1.8E-03		1.1E-02	0.1		0.1	0.2		0.2
Family Labor*Phosphate Use	-1.1E-02		8.1E-03	-0.1		0.1	-0.2		0.2
Nitrogen Use*Phosphate Use	-826.8	***	244.1	26.1	***	5.7	11.6	***	2.1
HybridDum*Acres	-0.3	***	0.1	-0.6	***	0.2	-0.7	***	0.2
HybridDum*Seed Quantity	70.7	***	11.3	8.6	***	1.8	5.2	***	1.1
HybridDum*Hired Labor	-8.0E-03	***	2.9E-03	-0.1	***	1.6E-02	-0.1	***	2.8E-02
HybridDum*Family Labor	-3.4E-04	***	7.7E-05	-1.9E-02	***	3.4E-03	-0.1	***	1.3E-02
HybridDum*Nitrogen Use	22.2	*	11.5	1.4		1.5	1.1		0.8
HybridDum*Phosphate Use	-5.6		6.8	-0.3		1.3	-0.5		0.7
Acres^2	-3.2E-02		3.2E-02						
Seed Quantity^2	109.9		1445.0						
Hired Labor^2	-4.2E-05		1.2E-04						
Family Labor^2	-1.4E-09		5.7E-08						
Nitrogen Use^2	-232.7	***	93.8						
Phosphate Use^2	1028.1	***	194.1						
Model Performance									
N	951			951			951		
R^2	0.60			0.63			0.63		
F Statistic	33.45	***		44.19	***		44.47	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

Table 5-4 shows the WLS regression results for the interaction effects specific to the quadratic and general Leontief functional forms. An F-test was conducted on all interaction effects to determine significance. The F statistic in the quadratic functional form is 26.96 and significant at the one percent level, while the F statistic for the Generalized Leontief forms is 32.53 and is also significant at the one percent level. From the data in Table 5-4, it is shown that all three forms show significance at the one percent level for the following interactions: Acres*Seed Quantity, Seed Quantity*Family Labor, Hired Labor*Family Labor, Nitrogen Use*Phosphate use, HybridDum*Acres, HybridDum*Seed Quantity, HybridDum*Hired Labor, and HybridDum*Family Labor. In terms of the squared variables, nitrogen use, which was negative, and phosphate use, which was positive, are the only squared term that exhibits significance.

Table 5-5. Variable Interaction Effects for Maize Plots

Positive	Negative
Acres*Hired Labor	Acres*Seed Quantity
Acres*Family Labor	Acres*Nitrogen Use
Acres*Phosphate Use	Seed Quantity*Hired Labor
Seed Quantity*Family Labor	Seed Quantity*Nitrogen Use
Seed Quantity*Nitrogen Use	Hired Labor*Nitrogen Use
Seed Quantity*Phosphate Use	HybridDum*Acres
Nitrogen Use*Phosphate Use	HybridDum*Phosphate Use
HybridDum*Nitrogen Use	
Hired Labor2	
Phosphate Use2	

Table 5-6. Marginal Productivity of Inputs for Maize Plots

	Linear		**Quadratic**		**Cobb-Douglas**		**Generalized Leontief (r=2)**		**Generalized Leontief (r=3)**	
Variable	MP		MP		MP		MP		MP	
Acres	-2.9E-02	***	-0.1		-0.3	***	-38.1	***	-54.0	***
Seed Quantity (tons per acre)	21.4	***	29.4	***	28.4	**	358.4	***	348.9	***
Hired Labor (days per acre)	3.4E-04		-2.9E-03		-2.7E-02		-0.7		-0.5	
Family Labor (hours per acre)	1.8E-05		-1.9E-04		-3.7E-04	***	-3.1E-02	*	0.0	
Nitrogen Use (tons per acre)	8.3	**	10.7		43.5	***	2434.6		1988.9	
Phosphate Use (tons per acre)	7.3	**	0.8		7.3		-1686.3		-1590.4	
Harvest Dummy	0.2	***	0.2	***	0.7	***	0.2	***	0.2	***
Hybrid Dummy	0.2	***	0.3	***	0.5	***	0.4	*	0.4	
AEZ2 Dummy	-0.1	***	-0.1	***	-0.5	***	-0.1	**	-0.1	**
AEZ4 Dummy	0.1	***	0.1	**	0.3	***	0.1	***	0.1	***
Terraced Dummy	-0.1	***	-0.1	***	-0.3	***	-0.1	**	-0.1	**
Credit Dummy	-1.4E-03		1.7E-02		-0.2	**	1.3E-02		1.3E-02	
Tenure Dummy	0.1	**	0.1	***	0.1		0.1	***	0.1	***
Years Hybrid Maize Experience	2.4E-03	*	2.6E-03	*	8.7E-03	**	2.8E-03	**	2.8E-03	**

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

5.3 Marginal Productivity Results

Marginal productivity effects for each input examine the change in output given one additional unit of input. Table 5-6 gives the marginal productivities for all inputs characterized by their functional forms. These marginal productivities were calculated by taking the derivative of each functional form with respect to each individual variable, $\partial f / \partial X_i$. The following are considered significant at the one percent level in the majority of the functional forms: acres, seed quantity, harvest dummy, hybrid dummy, AEZ2 dummy, AEZ4 dummy, terraced dummy, and tenure dummy.

Taking a closer look at the marginal product effects of each input, first, the results show that acres is negative in all forms, except the linear and Cobb-Douglas forms, indicating that for an increase in acreage, maize yields will increase. Seed quantity is also positive and significant across models, with an additional kilogram of seed increasing output approximately 20 to 30 kilograms per acre. The marginal effects of hired labor appear to be small, and in some forms, negative; however, the marginal effects for hired labor are not significant in any form. Family labor, interestingly, is only significant in the Cobb-Douglas and Generalized Leontief (r=2) form. In the Generalized Leontief form, family labor is small and positive, indicating that for an hour increase in labor per acre, output increases by only a small amount. This may indicate that family labor is being used at or near optimality, with families still trying to gain increased returns at a decreasing rate. The marginal effect of fertilizer use is not significant across the majority of models, but is positive in the linear forms, indicating that as an additional kilogram of fertilizer is applied per acre, output will increase by 7 to 8 kilograms. Last, the marginal effect of experience with hybrid maize seed is significant in all forms; however, the marginal effect is both positive and small. This indicates that increased experience with hybrid maize contributes only a very small amount towards increasing maize yields.

Next, looking at the marginal effects of the dummy variables all dummies, with exception to the credit dummy, are significant across the majority of models. First, the impact of the harvest dummy shows, on average, an increase of 0.2 tons per acre in yield when producing in the main season, which experiences longer rains. Next, looking at the hybrid dummy, the results show a small, positive increase in yield output for those using hybrid seed varieties. This further supports the argument that hybrid varieties boost average yields. The next two dummies for

AEZ2 and AEZ4 show interesting marginal effects for each zone. Those households who farm in AEZ2 actually experience a decrease in average yields by producing in this zone, while households who farm in AEZ4 experience a small increase in average yields. These dummies capture the diversity between these two zones, in which AEZ2 is more susceptible to severe drought conditions, while AEZ4 has the benefit of reliable rainfall. One surprising find within the dummy variables, is that of the terraced land dummy. The results show that those producing on terraced land actually experience a decrease in output, which contradicts the notion that terracing benefits farmers and can improve yields. This could be explained by the fact that despite efforts to make the best of drought conditions by terracing, this practice only benefits farmers if there is adequate rainfall needed for plant survival. In this sample, where the zones are highly susceptible to drought conditions, terracing may do more harm than good. Last, the tenure dummy, as expected, is positive and indicates that farmers who own their land experience slightly increased output. This could be for any number of reasons. For example, farm owners in Kenya are doing so to provide both food and additional income for their family. Since these farmers see the direct impact of their efforts on their livelihood, they are more likely to act in ways or adopt practices that impact output positively. The marginal productivities of each input provide information about the impact of the input upon maize yields in Kenya. This analysis can provide Kenyans with the information needed to choose an input mix that optimizes yield given their production choices and needs.

Table 5-7. OLS Regression Results for Variance Model

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error
INTERCEPT	-5.3	***	0.3	-5.0	***	0.7	-4.5	***	0.9	-6.6	***	1.3	-7.5	***	2.7
Acres	-0.1	*	0.1	-0.5		0.3	2.5E-03		0.1	0.5		0.7	0.8		1.6
Seed Quantity (tons/acre)	60.0	**	29.9	83.7		122.5	-0.2	**	0.1	45.6	***	17.7	31.5	**	15.0
Hired Labor (days/acre)	9.6E-03		1.2E-02	-2.5E-02		4.7E-02	0.1		0.1	-0.4		0.3	-1.8	***	0.7
Family Labor (hours/acre)	2.2E-04		3.1E-04	9.7E-04		9.5E-04	0.1		0.1	0.1	**	4.5E-02	0.5	**	0.2
Nitrogen Use (tons/acre)	0.1		13.4	-126.0	**	60.3	-0.4	***	0.2	1.2		16.6	38.9	*	20.0
Phosphate Use (tons/acre)	17.5		13.5	112.1	***	43.9	0.2	*	0.1	-4.9		17.2	-47.9	***	19.0
Harvest Dummy	0.7	***	0.2	0.5	***	0.2	-0.7	***	0.2	0.6	***	0.1	0.7	***	0.2
Hybrid Dummy	0.5	***	0.2	0.7		0.5	-0.4	***	0.2	1.0		1.1	1.5		1.7
AEZ2 Dummy	0.2		0.3	0.4		0.3	1.3	***	0.2	0.4		0.3	0.4	*	0.3
AEZ4 Dummy	0.8	***	0.2	0.7	***	0.2	0.4	*	0.2	0.5	**	0.2	0.6	**	0.3
Terraced Dummy	2.4E-02		0.2	0.3		0.2	0.3	*	0.2	0.3	*	0.2	0.4	**	0.2
Credit Dummy	0.1		0.2	0.1		0.2	0.9	***	0.2	0.2		0.2	0.3		0.2
Tenure Dummy	0.4	***	0.1	0.3	**	0.2	0.1		0.1	0.3	**	0.1	0.3	*	0.1
Years Hybrid Maize Experience	5.6E-03		8.1E-03	4.8E-04		8.1E-03	-1.3E-02		8.4E-03	7.1E-03		7.7E-03	7.7E-03		7.9E-03
Model Performance															
N	951			951			951			951			951		
R^2	0.13			0.15			0.13			0.18			0.21		
F Statistic	9.87	***		7.37	***		10.4	***		7.47	***		6.95	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

5.4 Variance Model Results

Table 5-7 gives the OLS regression results associated with all functional forms for the variance model. These results show that the following inputs are considered to be variance increasing amongst the linear model: seed quantity, nitrogen use, phosphate use, harvest season, hybrid seed, agro-ecological zone 2 location, agro-ecological zone 4 location, credit, and land tenure. On the other hand, acres is the only variable considered to be variance decreasing in the linear model. In the quadratic model, seed quantity, phosphate use, harvest season, hybrid seed, agro-ecological zone 2 location, agro-ecological zone 4 location, terraced land, credit, and land tenure are considered to be variance increasing inputs. Variance decreasing inputs in this model are acres and nitrogen use. Now, looking at the Cobb-Douglas model, it is clear that hired labor, family labor, phosphate use, agro-ecological zones 2 and 4 location, terraced land, credit, and land tenure are considered to be variance increasing inputs. On the other hand, seed quantity, nitrogen use, harvest season, and hybrid seed are considered to be risk decreasing. Last, looking at both generalized Leontief forms the results show that the following inputs are variance increasing: acres, seed quantity, family labor, nitrogen use, harvest season, hybrid seed, agro-ecological zones 2 and 4 location, terraced land, credit, and land tenure. In contrast, hired labor and phosphate use are considered to be variance decreasing inputs.

Breaking down the OLS regression results for the variance model by variable, this research looks at specific considerations. First, looking at the variables for harvest season, AEZ4, and land tenure, shows that these variables are consistent and significant across models. Harvest season is considered to be variance increasing, and indicates that for the main season, where long rains exist, there is a greater potential for increased variability. Next, AEZ4 is also positive across models indicating that fields within this region experience increased variance, as a direct result of different geographical factors. Last, tenure is positive and indicates that for land owned within members of a household or family members of that household, yields increase possibly as a result of an increased sense of ownership towards land and crops.

Next, examining variables that provide less consistency across models, acres is considered to be a variance decreasing input in the linear model, also the only significant form. In Kenya, this would be expected as more land is considered to be variance reducing. In this case, as land increases under favorable conditions, a producer is better off. However, the results show that field has very little impacts on variance, except in the linear form. The results show that larger fields are closer to the average. Seed quantity, as expected, is variance increasing, because additional seed results in higher yields. Next, looking at family labor, this variable is positive and small. This can be an indicator of households using family labor optimally at the field level. Interestingly, both nitrogen and phosphate rates are considered to be both variance reducing and increasing in different models. The last thing to be pointed out in this table is the hybrid varieties are only significant in the linear and Cobb-Douglas forms. This indicates that when controlling for all other factors in production, hybrid varieties are not a source of yield variability. This shows that though hybrids exist and have potential to increase yields, the yield increase alone is not enough to justify the costs of purchasing hybrid seed varieties.

Table 5-8. OLS Regression Results for Skewness Model

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error
INTERCEPT	-8.0	***	0.4	-7.5	***	1.0	-6.7	***	1.4	-9.9	***	2.0	-11.2	***	4.1
Acres	-0.2	*	0.1	-0.7		0.50	0.0		0.2	0.7		1.1	1.3		2.5
Seed Quantity (tons/acre)	90.0	**	44.9	125.6		183.7	-0.3	**	0.2	68.4	***	26.5	47.2	**	22.6
Hired Labor (days/acre)	1.4E-02		1.8E-02	-3.8E-02		0.1	0.1		0.1	-0.7		0.5	-2.8	***	1.1
Family Labor (hours/acre)	3.3E-04		4.7E-04	1.5E-03		1.4E-03	0.1		0.1	0.1	**	0.1	0.7	**	0.4
Nitrogen Use (tons/acre)	0.2		20.1	-189.0	**	90.4	-0.7	***	0.2	1.7		24.9	58.3	**	30.1
Phosphate Use (tons/acre)	26.3		20.2	168.2	***	65.8	0.4	*	0.2	-7.4		25.9	-71.9	***	28.6
Harvest Dummy	1.1	***	0.2	0.7	***	0.2	-1.0	***	0.2	0.9	***	0.2	1.0	***	0.2
Hybrid Dummy	0.7	***	0.3	1.1		0.7	-0.7	***	0.3	1.5		1.6	2.2		2.5
AEZ2 Dummy	0.3		0.4	0.6		0.5	1.9	***	0.4	0.5		0.4	0.7	*	0.4
AEZ4 Dummy	1.2	***	0.3	1.1	***	0.4	0.6	*	0.3	0.8	**	0.4	0.9	**	0.4
Terraced Dummy	0.0		0.2	0.4		0.3	0.5	*	0.3	0.4	*	0.3	0.5	**	0.3
Credit Dummy	0.1		0.3	0.1		0.3	1.3	***	0.2	0.3		0.3	0.4		0.3
Tenure Dummy	0.6	***	0.2	0.5	**	0.2	0.2		0.2	0.5	**	0.2	0.4	*	0.2
Years Hybrid Maize Experience	8.3E-03		1.2E-02	7.3E-04		1.2E-02	-1.9E-02		1.3E-02	1.1E-02		1.2E-02	1.2E-02		1.2E-02
Model Performance															
N	951			951			951			951			951		
R^2	0.13			0.15			0.13			0.18			0.21		
F Statistic	9.87	***		7.37	***		10.4	***		7.47	***		6.95	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

5.5 Full Order Moment Skewness Model Results

The regression analysis results are shown in Table 5-8. These results indicate that several inputs affect the skewness of yields. While there are many ways for Kenyans to positively affect skewness, the data indicates that producers have willingly and rapidly adopted hybrid seed varieties aimed at targeting drought conditions or pest problems with almost half of the sampled population adopting these varieties into their production practices. Producers in Kenya have also adopted other inputs, considered to be skewness increasing, such as fertilizer which adds nutrients to the plant. These changes are easily implemented into production practices, and their presence in the Kenyan agricultural input market indicates producers' desire to insure crop yield success. This would indicate that producers in Kenya exhibit downside risk aversion, and want to make changes that impact skewness positively, and also reduces exposure to downside risk, such as drought conditions.

Table 5-8 presents the OLS regression results associated with all functional forms for the skewness model. These results show that the following inputs are considered to be skewness increasing within the linear model: seed quantity, phosphate use, harvest season, hybrid seed, agro-ecological zones 2 and 4, credit, and land tenure. Acres is the only variable considered to be skewness decreasing in the linear functional form. Looking at the quadratic functional form, phosphate use, harvest season, agro-ecological zones 2 and 4, terraced land, credit, and land tenure are downside risk increasing, while acres and nitrogen use are downside risk decreasing. In the Cobb-Douglas functional form, hired labor, family labor, phosphate use, agro-ecological zones 2 and 4, terraced land, credit, and land tenure are risk increasing inputs. In contrast, seed quantity, nitrogen use, harvest season, and hybrid seed are risk decreasing inputs. Last, in the generalized Leontief functional forms, the following inputs are considered risk increasing: acres, seed quantity, family labor, nitrogen use, harvest season, hybrid seed, agro-ecological zones 2 and 4, terraced land, credit, and land tenure. On the other hand, hired labor and phosphate use are risk decreasing inputs.

Now, it is necessary to look at the results by each independent variable. First, looking at variables that are consistent across forms within the model, the results show that these include: seed quantity, harvest season, AEZ4, and tenure dummy. Seed quantity, as expected, is skewness increasing as additional seeding provides the opportunity for increased yield, and is mitigating risk against crop failure. The harvest season dummy has a positive impact on skewness and indicates that risk of crop failure is reduced in the main season which experiences longer rains. In addition, the AEZ4 dummy has a positive impact on skewness indicating that by producing in this agro-ecological zone producers are increasing their chance of observing yield gains. Last, land tenure, or producing on one's own land, positively influences skewness, also increasing the chance of observing yield grains. On the other hand, looking at some of the variables that are less consistent across models, we see that both acres and hybrid variety are significant in two or fewer forms. This indicates that neither variable are a good indicator of skewness when controlling for all other factors, a conclusion that will be discussed further in Section 5.7 with respect to the variable for hybrid varieties.

The full order moment of skewness regression provides important information about the impact each variable has on maize yields. However, it is important to note that the interpretation of this data is limited as both positive and negative residuals for each observation are aggregated into one category of asymmetry. By taking the absolute value of the residual, these results have been cleansed of quantitative terms and qualitative interpretation regarding negative and positive skewness. The next section describes the problems associated with looking at only asymmetry with respect to skewness, and provides an analysis of partial order moments for skewness.

5.6 Partial Order Moment Skewness Results

For the partial order moment skewness results, the residual from the OLS regression in the mean model, u_i, is taken, and cubed to estimate the skewness model. However, unlike the full order moment of skewness, this estimation does not take the absolute value of the residual, and instead splits the regression into positive and negative deviations from the mean. The OLS regression results of the partial order moments are presented in Table 5-9 and Table 5-10. Table 5-9 shows the results for partial order moments of positive skewness, while Table 5-10 shows the results for partial order moments of negative skewness.

The first thing to note about Table 5-9 and Table 5-10, is that by splitting the observations by positive and negative skewness, different independent variables become significant in each model. First, looking at Table 5-9 specifically, the positive partial order moments for skewness can be analyzed. Of the independent variables, harvest season, AEZ4, and tenure display significance across the most models. First, looking at the variable for harvest season, this variable has a positive impact on yield distribution and acts to push people positively farther away from the mean, increasing yields. Similarly, fields located in AEZ4, act to positively impact yield distributions, providing more yield gain. Last, land tenure, in this case, land owned by the producer or by a family member of the producer has a positive impact on

Table 5-9. OLS Regression for Positive Partial Order Moments

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error
INTERCEPT	-8.1		0.7	-7.9	***	1.9	-6.1	***	1.7	-9.1	**	4.0	-9.3		8.4
Acres	0.1		0.2	-0.4		0.98	0.1		0.2	0.7		2.0	2.6		4.6
Seed Quantity (tons/acre)	79.7		89.0	110.4		420.65	-0.2		0.2	60.8		56.7	29.5		48.9
Hired Labor (days/acre)	4.4E-02		3.2E-02	-1.5E-03		0.12	0.1		0.1	-1.0		0.9	-4.0	*	2.1
Family Labor (hours/acre)	1.2E-03		7.7E-04	3.4E-03		3.6E-03	5.0E-02		0.1	0.1		0.1	0.8		0.7
Nitrogen Use (tons/acre)	27.4		39.4	-69.5		186.80	-0.5	*	0.3	4.6		52.7	79.3	**	35.5
Phosphate Use (tons/acre)	10.0		36.3	170.6		150.62	0.2		0.3	8.7		49.7	-82.1	**	36.2
Harvest Dummy	0.8	*	0.4	0.5		0.45	-0.9	***	0.3	1.1	***	0.4	1.0	***	0.4
Hybrid Dummy	0.8		0.5	0.9		1.54	-0.5		0.3	-0.8		2.8	-3.4		4.0
AEZ2 Dummy	0.4		0.7	0.3		0.73	2.0	***	0.4	1.1		0.7	1.1		0.7
AEZ4 Dummy	1.3	**	0.6	1.1		0.71	0.8	**	0.4	1.3	**	0.6	1.4	**	0.6
Terraced Dummy	-0.1		0.5	1.9E-02		0.51	0.4		0.3	-1.7E-02		0.5	0.2		0.5
Credit Dummy	-0.3		0.5	0.3		0.53	1.5	***	0.3	0.3		0.5	1.0	**	0.5
Tenure Dummy	0.8	**	0.4	0.4		0.44	0.1		0.3	0.5		0.4	-0.1		0.4
Years Hybrid Maize Experience	-2.6E-03		2.1E-02	2.4E-02		0.02	-3.0E-02	**	1.4E-02	8.4E-03		2.1E-02	3.5E-03		2.2E-02
Model Performance															
N	396			409			544			407			399		
R^2	0.12			0.16			0.14			0.16			0.20		
F Statistic	3.59	***		1.72	***		6.0	***		1.97	***		2.63	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

Table 5-10. OLS Regression Results for Negative Partial Order Moments

	Linear			Quadratic			Cobb-Douglas			Generalized Leontief (r=2)			Generalized Leontief (r=3)		
	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error	Coeff		Std. Error
INTERCEPT	-8.0	***	0.5	-7.4	***	1.2	-7.1	***	2.6	-10.1	***	2.7	-9.7	*	5.9
Acres	-0.5	***	0.1	-0.9		0.6	-0.1		0.3	0.5		1.3	-0.7		3.2
Seed Quantity (tons/acre)	132.8	**	55.0	182.8		262.2	-0.5		0.3	64.4	*	35.3	43.3		32.7
Hired Labor (days/acre)	-9.5E-03		2.1E-02	-4.8E-02		0.1	0.2		0.3	-0.6		0.6	-2.5	*	1.5
Family Labor (hours/acre)	-3.9E-04		5.5E-04	4.9E-04		2.0E-03	0.2		0.2	0.1		0.1	0.2		0.6
Nitrogen Use (tons/acre)	-10.2		19.6	-295.8	*	162.4	-0.9	**	0.4	-56.7		43.1	17.3		47.9
Phosphate Use (tons/acre)	30.8		21.5	260.2	**	120.2	0.6		0.4	29.3		40.6	-39.4		44.4
Harvest Dummy	1.2	***	0.3	0.9	***	0.3	-1.4	***	0.4	0.8	***	0.3	1.0	***	0.3
Hybrid Dummy	0.9	***	0.3	1.3		1.0	-1.0	**	0.5	4.3	**	2.0	7.8	***	3.1
AEZ2 Dummy	0.2		0.5	0.7		0.6	2.1	***	0.8	0.2		0.5	0.2		0.5
AEZ4 Dummy	1.0	***	0.4	0.9	*	0.5	0.3		0.6	0.6		0.4	0.7		0.5
Terraced Dummy	0.1		0.3	0.5	*	0.3	0.5		0.4	0.9	***	0.3	0.7	**	0.3
Credit Dummy	0.5	*	0.3	2.7E-02		0.3	1.0	**	0.5	0.3		0.3	0.3		0.3
Tenure Dummy	0.6	**	0.3	0.6	**	0.3	0.3		0.4	0.6	**	0.3	0.8	***	0.3
Years Hybrid Maize Experience	1.6E-02		1.4E-02	-1.7E-02		0.0	-5.4E-03		2.0E-02	7.7E-03		1.4E-02	1.3E-02		1.5E-02
Model Performance															
N	555			542			407			544			552		
R^2	0.17			0.16			0.14			0.24			0.26		
F Statistic	7.86	***		2.79	***		4.5	***		4.58	***		5.31	***	

***,**,* indicates significance at the 1%, 5% and 10% levels respectively.

skewness. One variable that is unexpectedly missing from this category is the variable for hybrid varieties. This variable was found to be insignificant across all models and, therefore, does not have a positive impact upon skewness.

Next, looking at Table 5-10 the results of the OLS regression for negative partial order moments of skewness are given. First, harvest season is also significant in the negative partial order moments model across all models. This indicates that while harvest can positively impact skewness, there is also the potential for the main season to provide low yield potential. This could be for any number of reasons, such as reliance on rainfall in this season for crop production. Another explanation might be that during the main season where longer rainfalls are experience, individuals think they have longer amounts of time and, as a result, do not produce in a timely fashion. Like the positive partial order model, the AEZ4 variable is also significant, indicating, again, the idea that observations within this zone, while having increased yield potential, also have more to lose in unfavorable conditions. Land tenure is also significant across a number of models, and indicates that this variable has a negative impact on yield distribution and acts to push people negatively farther away from the mean, decreasing yields. Last, the variable for hybrid variety provides interesting contrasts between the positive and negative partial order moment models. This variable is significant across four forms and indicates that hybrids have a positive impact on yield distribution and acts to push people closer to the mean. The results indicate that hybrid varieties increase yields, and allow producers who experience yields lower than average to move closer to the mean.

The important point to note after looking at both the full order moment of skewness, as well as the partial order moments, is that each models provides different results. The most obvious way that the models do this is by different significances for the same variables in separate models. To test the hypothesis of symmetric effects of inputs, a Chow test is used to test for equality of the parameters in equations 7 and 8. The Chow test is F distributed (14, 14) and the critical value is 2.4 at the 5 percent level. The results of the Chow test statistic are 1.25 in the linear form, 0.61 in the quadratic, 0.7 in the Cobb-Douglas, 0.7 in the generalized Leontief (r=2), and 0.9 in the generalized Leontief (r=3). These values mean that we fail to reject the null hypothesis that coefficients are the same within both the positive and negative partial order moment results for skewness.

5.7 Hybrid versus Open Pollinated/Traditional Varieties

The hybrid versus local seed varieties debate has been a subject of dispute since the United States began to market hybrid seed varieties in the market. Hybrid varieties are seemingly the solution to all production problems, and are argued to perform better than OPVs in water stressed areas. If this is the case, then why have all producers not adopted these varieties? As mentioned previously, many producers associate a diminished productivity associated with hybrid varieties in drought conditions. In this research, among the sample, over half of the population still plants traditional varieties over hybrid varieties. However, it should be noted that the remainder of the population, does in fact plant hybrid varieties, which indicates a growing acceptance towards the productivity of these varieties. This research contributes to the debate, and provides a clearer picture of the situation in Kenya, specifically, within the maize plots examined in this research.

Mean, Variance, and Skewness

The mean, variance, and skewness models show that in all models, hybrid varieties are increasing in average yield, variance, and positively affect skewness in the partial order moments. In the mean model, the presence of hybrid seed increases average yields by anywhere from 0.2 to 0.5 tons per acre. The variance model shows that hybrid seed varieties are variance increasing in the linear form, indicating more variability and less stability for producers. Looking at the impact of hybrid seed varieties on yield, this means that hybrids have a greater potential to increase yields in favorable conditions. However, this also means that in unfavorable conditions, the hybrid seed varieties have the potential to perform as bad as or worse than open pollinated or traditional varieties. Since average yield increases with hybrid seed in this data set, despite the presence of drought conditions in 2007, it appears that hybrid varieties are living up to their potential. Returning the conclusion that hybrid seed varieties are variance increasing, this concept contradicts the popular belief that hybrid varieties are variance decreasing.

Breaking the skewness model into positive and negative partial order moments, the results show that hybrid varieties are not significant in the positive partial order moment model and, therefore, do not impact positive deviation from the mean, and have a positive impact on yield distribution in the negative partial order moment model. While hybrid seed varieties do not reduce variance, they are considered to be skewness increasing in the full order moment of the model. Although seemingly the same concept, there is a clear discrepancy in the nomenclature and it should be noted that risk reducing is not equivalent to variance decreasing in the variance model. This means that hybrid varieties cause the asymmetry or skewness of the distribution to be distributed further to the right. However, when the skewness model is evaluated in terms of positive and negative partial order moments, hybrid varieties are found to have a positive impact on skewness. Specifically, for individuals whose mean yield was less than average, those using hybrid pulled those closer to the mean. In the case of Kenya, hybrid varieties are increasing in importance and it is important for both producers and consumers to understand that these varieties are variance increasing, contrary to popular belief, indicating that yield variability is increased.

This research shows the impact of hybrid seed varieties in the production practices of Kenyan agricultural producers. While there is no clear answer, from this research, on whether hybrid or traditional varieties are the best in Kenya, it should be understood that hybrids, despite their ability to target specific components of downside risk, have not and will not solve Kenya’s food production problems overnight. Current hybrids may still be unprofitable under water stress. Many other factors, such as severe drought conditions, are contributing to the increased probability of crop failure for farmers, and as a result, other solutions, in addition to the creation of hybrid varieties should be sought out. It is also important, from a policy stand point to identify ways to provide hybrid seed varieties at a lower price, so that producers can better justify hybrid purchasing in order to increase yields.

Chapter 6 - Conclusion

The objective of this research was to determine if there is variability in Kenyan maize yields, and whether or not specific inputs, specifically hybrid varieties, are both variance and skewness increasing or decreasing within the data set. This research gives an empirical analysis of the effect of various inputs on farm productivity. In 2010, the Bill and Melinda Gates Foundation provided funds committed to the development of hybrid varieties aimed at small farm productivity, as well as reducing exposure to downside risk, or the probability of crop failure for agricultural producers in arid environments (Gates Foundation, 2010). This research was designed to help give a better understanding of how a number of inputs, including seed varieties, impact agricultural production, and how resources from organizations, like the Gates Foundation, should be distributed.

The results from this research show that hybrid seed varieties can increase farm productivity, but do not reduce yield variability for Kenyan maize producers. The empirical analysis shows that hybrid varieties have a crucial role to play in promoting increased maize yields. The results of this research prove that hybrid varieties are not, contrary to popular belief, variance reducing, but are, in fact, variance increasing. Hybrid varieties positively impact skewness in the full order of moments, and, as a result reduce downside risk exposure, or the probability of crop failure. When partial order moments of skewness are analyzed, hybrid varieties are revealed to have a positive impact on skewness, pulling producers toward the mean yield. This is a good result when looking at positive skewness, but bad when looking at negative skewness. Although hybrid varieties can provide increased yields, they do not protect producers from yield variability and may expose them to increased downside risk exposure. While hybrid varieties are an easy adaptation to increase average yields, they are not the solution to all of Kenya’s agricultural problems. Currently, one of Kenya’s biggest problems is that it relies heavily on rainfall for agricultural production, and in years of severe drought, crops, hybrid or not, have no chance at survival. Organizations like the Gates Foundation and the Rockefeller Foundation should continue to focus efforts on producing Kenyan drought tolerant maize varieties, but should also focus on creating access to other inputs, such as irrigation when aiming to boost agricultural yields.

This analysis provided for analysis of stochastic factors by the inclusion of dummy variables targeted at agro-ecological conditions and land quality in specific geographical locations and on a field level basis. The presence of two dummies for agro-ecological zones 2 and 4 identifies the positive impact land in zone four and the negative impact that land in zone two has on average maize yields. This was expected as zone two has locations where rainfall is distributed unequally, with Mount Kenya preventing rain-bearing winds from reaching producers in this zone. On the other hand, locations in zone four are often wet most of the year while varying between hot and cold temperatures in different districts.

Other interesting results of this research are the impact that labor and seed quantity has upon farm productivity and yield variability for producers. Hired labor has very little influence on mean, variance, and skewness of maize yields. Family labor has a slightly greater influence on these factors, but both variables are among the lowest in terms of boosting mean yield, variance, or skewness. This could be explained by the fixed nature of labor in Kenya, with family labor being limited, to a certain extent, and hired labor being constrained by income. Next, looking at seed quantity, this research shows that additional tons of seed per acre have the greatest impact on boosting mean yields for Kenyan agricultural producers. This conclusion shows that there are likely seed productivity gains to be made by modifying planting density among producers. Efficiency gains will have to be made up until producers are producing in the second stage of production where yields are increasing at a decreasing rate until maximum efficiency is reached. After this point, producers will begin to experience decreasing yields as additional seed results in crowding of plants to the point where yield is lost.

In terms of overall impact of each variable on mean, variance, and skewness of maize yields seed quantity, nitrogen use, and hybrid seed contribute the most to influencing these factors. In contrast, years of experience with hybrid maize, land tenure, terraced land and labor have the least influence on mean, variance and skewness within this research. This research, specifically, knowing which inputs provide the greatest or least gains, provides organizations like the Gates Foundation and the Rockefeller Foundation with the knowledge to determine the best way to improve smallholder productivity in Kenya. In recent years, hybrid seed varieties have gained popularity, and have the potential to solve crucial production problems in Kenya. If these varieties are paired with the knowledge of how to maximize yield with specific inputs, Kenya has the opportunity to see substantial productivity gains throughout the country, especially in arid and semi-arid regions.

The research presented in this paper is one of many additions to the analysis of the Kenyan data set provided by Tegemeo Institute of Agricultural Policy and Development and Michigan State University. Research has been completed analyzing policy implications, fertilizer use, and production trends; additionally, this research provides an analysis of the impact hybrid varieties have on the mean, variance and skewness of yield. There is potential for further research to be completed with the data set analyzing production as a whole or, additionally, looking at components of production in Kenya, such as the impact of family labor upon productivity. This research could also benefit from an analysis of net returns, which include both revenue and cost figures. Such research would go beyond the scope of this research, which provides solely a production model. In addition, there is also potential for panel data analysis of this data as the data set provides multiple household observations over several years. This analysis would provide a better picture of how Kenyan agricultural producers are changing both their inputs and outputs over time to meet needs.

References

Anderson, J. R., & Hazell, P. 1989. *Variability in Grain Yields*. London: The John Hopkins University Press.

Antle, J. 2010. "Asymmetry, Partial Moments, and Production Risk." *American Journal of Agricultural Economics* 92(5): 1294-1309.

Ariga, J., Jayne, T. S., Kibaara, B., and Nyoro, J. 2008. "Trends and Patterns in Fertilizer Use by Smallholder Farmers in Kenya, 1997-2007." Working paper, Nairobi: Tegemeo Institute of Agricultural Policy and Development.

Babcock, B. A. 1995. "The Cost of Agricultural Production Risk." *Agricultural Economics* 12(2), 141-150.

Battese, G. E. 1997. "A Stochastic Frontier Production Function with Flexible Risk Properties." *Journal of Productivity Analysis* 8(3), 269-80.

Beattie, B.R., Taylor, C.R., and Watts, M.J. 2009. *The Economics of Production.* 2nd ed. Malabar, FL: Krieger Publishing.

Bett, C. 2004. "Farm Level Adoption Decisions of Soil and Water Management Technologies in Semi-arid Eastern Kenya." Paper presented at Australian Agricultural and Resource Economics Conference, Melbourne, Australia 11-13 February.

Binswanger , H. 1981. Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India. *Economic Journal* 91(364), 867.

Bullock, D. G. 1994. Quadratic and Quadratic-plus-plateau Models for Predicting Optimal Nitrogen Rate of Corn: A Comparison. *Agronomy Journal* 86(1), 191-195.

Chamberlin, J., and Jayne, T. S. 2009. "Has Kenyan Farmers' Access to Markets and Services Improved? Panel Survey Evidence, 1997-2007." Working paper, Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development.

CIMMYT. (2010). *About Us.* Retrieved 9/15/2010, http://www.cimmyt.org/index.php/en/about-us.

Critchley, W. 1991. *Looking After Our Land: Soil and Water Conservation in Dryland Africa*. Oxford: Oxfam Print Unit.

D'Agostino, R. B., Balanger, A. and D'Agostino, R. B., Jr. 1990. "A Suggestion for Using Powerful and Informative Tests of Normality." *The American Statistician* 44: 316-321.

Di Falco, S and Chavas, J. 2006. "Crop Genetic Diversity, Farm Productivity and the Management of Environmental Risk in Rainfed Agriculture." *European Review of Agricultural Economics* 33(3), 289.

Doss, C., McPeak, J., and Barrett, C. 2005. "Perceptions of Risk Within Pastoralist Households in Northern Kenya and Southern Ethiopia." Paper presented at AAEA annual meeting, Providence, RI, 24-27 July.

Edmeades, G. O. 2008. Drought Tolerance in Maize: An Emerging Reality. A Feature in James, Clive. 2008. Global Status of Commercialized Biotech/GM Crops: 2008. *ISAAA Brief* No. 39. ISAAA: Ithaca, NY.

Food and Agriculture Organization. (2010). *FAOSTAT Commodities by Country.* Retrieved 9/29/2010, 2010, from http://faostat.fao.org/site/339/default.aspx.

Gardner, B., and Rausser, G. 2001. *Handbook of Agricultural Economics*. Amsterdam, The Netherlands: Elsevier Science B.V.

Gates Foundation. (2010). *Our work in agricultural development - Overview and approach.*, 2010, from http://www.gatesfoundation.org/agriculturaldevelopment/Pages/overview.aspx#.

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

Griffiths, W. 1982. "Using Time-series and Cross-section Data to Estimate a Production Function with Positive and Negative Marginal Risks." *Journal of the American Statistical Associatio,* 77(379), 529-36.

Isik, M. 2003. "Stochastic Technology, Risk Preferences, and Adoption of Site-specific Technologies." *American Journal of Agricultural Economics* 85(2), 305.

Isik, M. 2002. "Resource Management Under Production and Output Price Uncertainty: Implications for Environmental Policy." *American Journal of Agricultural Economics* 84(3), 557-71.

Jia, X. 2009. "Synergistic Green and White Revolution: Evidence from Kenya and Uganda." Paper presented at International Association of Agricultural Economists Conference, Beijing, China, 16-22 August.

Just, R. E and Pope, R.D. 1978. "Stochastic Specification of Production Functions and Economic Implications." *Journal of Econometrics* 7 (1978): 67-86.

Just, R. E. and Pope, R.D. 2003. "Agricultural Risk Analysis: Adequacy of Models, Data, and Issues." *American Journal of Agricultural Economics* 85(5): 1249-1256.

Karanja, D. D. 2003. "Welfare effects of maize technologies in marginal and high potential regions of Kenya." *Agricultural Economics* 29(3), 331.

Kimenju, S., and Tschirley, D. 2008. "*Agriculture and Livelihood Diversification in Kenyan Rural Households*." Working paper, Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development.

Kumbhakar, S. C. 2010. "Estimation of Production Risk and Risk Preference Function: A Nonparametric Approach." *Annals of Operations Researc*, 176(1): 369.

Langyintuo, A. S., Mwangi, W., Diallo, A. O., MacRobert, J., Dixon, J., & Banziger, M. 2008. *An Analysis of the Bottlenecks Affecting the Production and Deployment of Maize Seed in Eastern and Southern Africa*. Harare, Zimbabwe: CIMMYT.

Library of Congress. 2007. *Country Profile: Kenya.* Library of Congress Federal Research Division.

Lybbert, T., and Bell, A. 2010. "Stochastic Benefit Streams, Learning, and Technology Diffusion: Why Drought Tolerance is Not the New Bt." *AgBioForum* 13(1): 19-24.

Mabaya, E., Cramer, L., Mahiga, V., Pham, H., Simpson, T., and Tang, X. 2009. "Supplying Improved Seed to Farmers in Rural Kenya: The Case of Freshco Kenya ltd." *International Food and Agribusiness Management Review* 12(1).

Monsanto. 2010a. *Monsanto at a Glance.* Retrieved 9/10/2010, 2010, from http://www.monsanto.com/whoweare/Pages/default.aspx.

Monsanto. 2010b. *Monsanto Product Overview.* Retrieved 9/10/2010, 2010, from http://www.monsanto.com/products/Pages/default.aspx.

Muhammad, L., Mwabu, D., Mulwa, R., and Mwangi, W. 2008. "*DTMA Household Survey Report: Kenya."* Nairobi, Kenya: Kenya Agricultural Research Institute and International Maize and Wheat Improvement Center (CIMMYT).

Naeem, S. 1994. ”Declining Biodiversity Can Alter the Performance of Ecosystems." *Nature* 368 (6473): 734-737.

Nyabiage, J. S. 2009. "Agrochemical Body says, GMOs Won't Drive Us Out." *Daily Nation.*

Nyoro, J. 2002. "Kenya's Competitiveness in Domestic Maize Production: Implications for Food Security. Working paper, Netherlands, University of Leiden.

Nyoro, J., Kirimi, L., and Jayne, T. S. 2004. "Competitiveness of Kenyan and Ugandan Maize Production: Challenges for the Future." Working paper, Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development.

Oehmke, J., Jayne, T. S., Aralas, S., and Mathenge, M. 2010. "Impacts of USAID/Kenya supported agricultural productivity interventions on household income and poverty reduction." Working paper, Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development.

Pioneer Hi-Bred International, I. 2010. *History.* Retrieved 9/29/2010, 2010, from http://www.pioneer.com/web/site/portal/menuitem.cc20eec90551c318bc0c0a03d10093a0/.

Pope, R and Kramer, R. 1979. ”Production Uncertainty and Factor Demands for the Competitive Firm." *Southern Economic Journal* 46(2): 489-501.

Primal Seeds. 2000. *Hybrid versus open pollinated Seeds.* 2010, from http://www.primalseeds.org/hybrid.htm.

Ramaswami, B. 1992. "Production Risk and Optimal Input Decisions." *American Journal of Agricultural Economics* 74(4): 860-869.

Shankar, B, Bennett, R., and Morse, S. 2007. "Output Risk Aspects of Genetically Modified Crop Technology in South Africa." *Economics of Innovation and New Technology* 16(4), 277-291.

Thurow, R. 2008. "Agriculture's Last Frontier." *The Wall Street Journal,* pp. A14.

Traxler, G, Falck-Zepeda, J., Ortiz-Monasterio, J.I., and Sayre, K. 1995. "Production Risk and the Evolution of Varietal Technology." *American Journal of Agricultural Economics 77*(1): 1-7.

Appendix A - Specification Tests

Table 0-1. F Test Results for Interaction Effects

Functional Form	F	Block Degrees of Freedom	Residual Degrees of Freedom	Pr > F	R2
Quadratic	26.96	27	923	0.00	0.4409
GL2	32.53	21	929	0.00	0.4238
GL3	32.53	21	929	0.00	0.4238

Appendix B - Pearson Correlation

Table 0-2. Pearson Correlation Coefficient Matrix for All Variables

	Variable	Yield	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
Yield	Pearson Correlation	1	-.154	.103	-.031	.136	.521	.520	.290	.449	-.082	.458	-.066	-.116	.042	.322
	Sig. (2-tailed)		.000	.001	.345	.000	.000	.000	.000	.000	.011	.000	.041	.000	.197	.000
X1	Pearson Correlation	-.154	1	-.271	-.136	-.363	.009	.011	.043	-.072	-.098	-.107	.343	.078	-.037	.040
	Sig. (2-tailed)	.000		.000	.000	.000	.774	.735	.182	.025	.003	.001	.000	.016	.260	.221
X2	Pearson Correlation	.103	-.271	1	.236	.209	-.046	-.004	-.034	-.083	.105	-.002	-.058	-.042	-.009	.016
	Sig. (2-tailed)	.001	.000		.000	.000	.155	.891	.288	.010	.001	.946	.072	.195	.793	.621
X3	Pearson Correlation	-.031	-.136	.236	1	-.099	-.069	-.063	-.084	.037	.238	-.176	.036	.122	.093	-.011
	Sig. (2-tailed)	.345	.000	.000		.002	.033	.053	.009	.254	.000	.000	.274	.000	.004	.739
X4	Pearson Correlation	.136	-.363	.209	-.099	1	.062	.091	-.064	.080	.090	.247	-.124	-.106	-.014	.088
	Sig. (2-tailed)	.000	.000	.000	.002		.058	.005	.048	.013	.006	.000	.000	.001	.659	.006
X5	Pearson Correlation	.521	.009	-.046	-.069	.062	1	.817	.190	.376	-.103	.514	.072	-.099	-.031	.340
	Sig. (2-tailed)	.000	.774	.155	.033	.058		.000	.000	.000	.001	.000	.026	.002	.344	.000
X6	Pearson Correlation	.520	.011	-.004	-.063	.091	.817	1	.182	.378	-.090	.516	.081	-.107	-.056	.380
	Sig. (2-tailed)	.000	.735	.891	.053	.005	.000		.000	.000	.006	.000	.013	.001	.082	.000
X7	Pearson Correlation	.290	.043	-.034	-.084	-.064	.190	.182	1	.219	-.050	.116	-.001	-.041	-.012	.119
	Sig. (2-tailed)	.000	.182	.288	.009	.048	.000	.000		.000	.122	.000	.984	.207	.713	.000
X8	Pearson Correlation	.449	-.072	-.083	.037	.080	.376	.378	.219	1	.077	.357	.023	-.049	.051	.392
	Sig. (2-tailed)	.000	.025	.010	.254	.013	.000	.000	.000		.018	.000	.475	.128	.114	.000
X9	Pearson Correlation	-.082	-.098	.105	.238	.090	-.103	-.090	-.050	.077	1	-.212	-.266	.001	.157	.137
	Sig. (2-tailed)	.011	.003	.001	.000	.006	.001	.006	.122	.018		.000	.000	.978	.000	.000
X10	Pearson Correlation	.458	-.107	-.002	-.176	.247	.514	.516	.116	.357	-.212	1	.021	-.228	-.039	.384
	Sig. (2-tailed)	.000	.001	.946	.000	.000	.000	.000	.000	.000	.000		.520	.000	.232	.000
X11	Pearson Correlation	-.066	.343	-.058	.036	-.124	.072	.081	-.001	.023	-.266	.021	1	.167	-.036	.070
	Sig. (2-tailed)	.041	.000	.072	.274	.000	.026	.013	.984	.475	.000	.520		.000	.272	.031
X12	Pearson Correlation	-.116	.078	-.042	.122	-.106	-.099	-.107	-.041	-.049	.001	-.228	.167	1	.015	-.085
	Sig. (2-tailed)	.000	.016	.195	.000	.001	.002	.001	.207	.128	.978	.000	.000		.650	.009
X13	Pearson Correlation	.042	-.037	-.009	.093	-.014	-.031	-.056	-.012	.051	.157	-.039	-.036	.015	1	.031
	Sig. (2-tailed)	.197	.260	.793	.004	.659	.344	.082	.713	.114	.000	.232	.272	.650		.340
X14	Pearson Correlation	.322	.040	.016	-.011	.088	.340	.380	.119	.392	.137	.384	.070	-.085	.031	1
	Sig. (2-tailed)	.000	.221	.621	.739	.006	.000	.000	.000	.000	.000	.000	.031	.009	.340	

X1 - Acres, X2 - Seed Quantity, X3 - Hired Labor, X4- Family Labor, X5 - Nitrogen Use, X6 - Phosphate Use, X7 - Harvest Dummy, X8 - Hybrid Dummy, X9 - AEZ2 Dummy, X10 - AEZ4 Dummy, X11 - Terraced Dummy, X12 - Credit Dummy, X13 - Tenure Dummy, X14 - Years Hybrid Maize Experience

A pairwise correlation between the variables in the linear functional forms was also tested to identify whether any variables were highly correlated with one another, and therefore, resulting in a lack of endogeneity in the model. The results of this correlation are shown in Table 0-2. The only highly correlated variables are nitrogen and phosphate, with a Pearson correlation coefficient of .817. This is a relationship that was expected within the data set and a result of fertilizers being broken down by element for this research. The remaining relationships between variables are, for the most part, close to zero, and therefore, are independent of one another. Table 0-2 also shows which pairwise correlation relationships are significant at the one, five, and ten percent level.