IMPACT OF STAPLE PRICE CHANGES ON SUPPLY RESPONSE OF MAIZE PRODUCTION: AN ANALYSIS OF HOUSEHOLD PANEL DATA IN ZAMBIA

By

Dingiswayo Ugwunkwo Chinwendu Banda

A THESIS

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

MASTER OF SCIENCE

Agricultural, Food and Resource Economics

2012

ABSTRACT

IMPACT OF STAPLE PRICE CHANGES ON SUPPLY RESPONSE OF MAIZE PRODUCTION: AN ANALYSIS OF HOUSEHOLD PANEL DATA IN ZAMBIA

By

Dingiswayo Ugwunkwo Chinwendu Banda

Global fluctuations in cereal prices since 2008 have created significant uncertainty and flux in international commodity markets. Projections are that prices over the next 10 years will generally be higher than they have been over the past half-century. Because of major heterogeneity in resources and farming conditions, farmers in agrarian societies like Zambia face a diverse range of challenges and opportunities in responding to these higher commodity prices. This thesis investigates the supply response of maize growing households in Zambia. A production function is created to identify the main determinants of ability to expand area under maize and maize yields among smallholder households.

The study uses panel survey data on 5,400 farm households from the 2001, 2004 and 2008 Supplemental survey to the Post Harvest Survey of 1999/2000 agricultural season. Fixed effects analysis is used to model the response of households to different explanatory variables such as maize prices, household demographic characteristics, and asset holdings.

I find that farm households' ability to respond to higher maize prices by expanding area under cultivation and yield are significantly affected by their resource endowment

DEDICATION

This thesis is dedicated to my dear family, Kaunda, Kondwani, Freda, Lindila and Daniel. Thank you for your love, support and continued assistance.

ACKNOWLEDGEMENTS

I wish to acknowledge the support of Dr. Thom Jayne, Dr. Gelson Tembo, Dr. Mike Webber, Dr. Jack Meyer, Dr. Jones Govereh, Dr. Nicky Mason, Dr. William Burke, Mr. Julius Shawa, Ms. Emma Malawo, Mr. Michael Isimwaa and the entire NEWU team for their support and assistance. I also wish to acknowledge the support of the entire IAPRI team.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMNS

AE	Adult Equivalent
CFS	Crop Forecast Survey
CSO	Central Statistical Office
FAO	Food and Agriculture Organisation
FBS	Food Balance Sheet
FRA	Food Reserve Agency
FSP	Fertilizer Support Programme
GDP	Gross Domestic Product
GRZ	Government of the Republic of Zambia
MACO	Ministry of Agriculture and Cooperatives
MOFNP	Ministry of Finance and National Planning
OECD	Organization of Economic Co-operation and Development
OLS	Ordinary Least Squares
PHS	Post Harvest Survey
PCA	Principal Components Analysis
SEA	Standard Enumeration Area
CSA	Census Supervisory Area
SS	Supplemental Survey
USDA	United States Department of Agriculture
USGS	United States Geological Survey

WRSI	Water Requirements Satisfaction Index
ZNFU	Zambia National Farmers Union

Introduction

Background

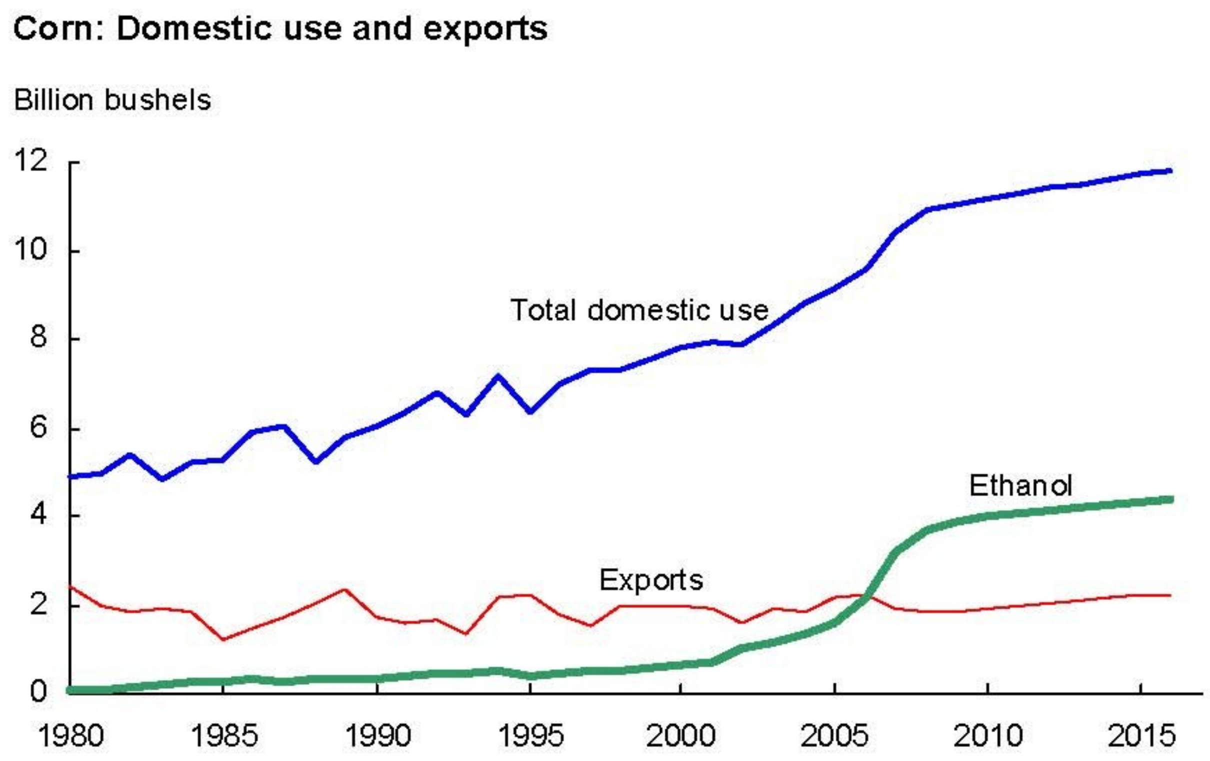
World grain prices are projected to rise over the next few years mostly due to the increasing demand for ethanol in large markets such as the USA and China. “In America, the annual rise in the producer-price index for finished consumer foods has picked up from a little over 1.5% last year to almost 4%.... Merrill Lynch has also coined an eye-catching term for the process: agflation (Economist, 2007). Most of this rise in demand for corn is for the production of ethanol.

The U.S. Department of Agriculture projects that world grain use will grow by 20 million tonnes in 2006. Of this, 14 million tons will be used to produce fuel for cars in the United States, leaving only 6 million tons to satisfy the world's growing food needs (USDA, 2007). Indeed according to the GAIA foundation, “In 2006, an increase in the use of grain worldwide for conversion to bio-fuels led to a 60% increase in global grain prices and speculator interest.” Estimates vary on the precise magnitude of the rise in prices but there is a consensus that, because of a US energy policy based on promoting the use of ethanol from corn, there will be upward pressure on world grain prices. This secular rise in US grain prices is likely to be transmitted globally, to some extent even into landlocked African countries. Strong demand for ethanol production results in higher corn prices and provides incentives to increase corn acreage. Much of this increase occurs by adjusting crop rotations between corn and soybeans, causing a decline in soybean plantings. As the ethanol industry absorbs a larger share of the corn crop, higher prices will affect both domestic uses and exports, providing for more intense competition between and among the domestic industries and foreign buyers in the demand for feed grains. U.S. feed use of corn typically accounts for 50-60 percent of total corn use and the United States typically accounts for 60-70 percent of world corn exports. Market adjustments to higher prices result in a reduced share of corn used directly for domestic livestock feeding and a lower U.S. share of global corn trade (USDA).[1]

Based on the projections by the USDA, fuel alcohol use of corn in the USA will rise faster than exports of corn.

[1] USDA, Long term projections, 2007

Figure 1: US corn production and use projections 1980 to 2015



USDA Long-term Projections, February 2007
Note: For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis

The prospects that world cereal prices will be higher than they have been in the recent past raises important questions about the ability of African farmers to respond to these incentives. Especially important is the distributional effects of higher food prices given the great heterogeneity among rural farm households. The smallholder[2] sector in Zambia typically comprises households that grow food for own consumption. Households with surplus production sometimes participate in sales. Many of these households face constraints of labour, land, assets etc. These constraints may or may not limit the ability of some households to participate in the market. Effective analysis of the impact of price changes on the small holder sector requires an understanding of the structure of the agriculture sector in Zambia.

Zambia has over 1,400,000 smallholder farmers[3] and approximately 1500 large scale farmers. Maize is the most widely grown crop in Zambia constituting over 48% of total area planted to the 19 major crops grown in the country during the 2008/09 season. Over the last 10 years, area under maize has typically ranged between 47-50% of total area under crops. A significant proportion of the maize grown in Zambia is produced by small holder farmers. During the 2008/09 season 88% of the 1,880,000 metric tons (mt) maize production was from the small holder sector[4]. Although maize production has tended to fluctuate significantly, due in large part to unpredictable weather

[2] A smallholder is formally defined as a small and medium scale farming household that cultivates a maximum of 19.99 hectares of crops and/or households raising 50 or more cattle, 20 or more pigs, 30 or more goats, and/or 50 or more chickens, even if they do not qualify basing on area under crops.

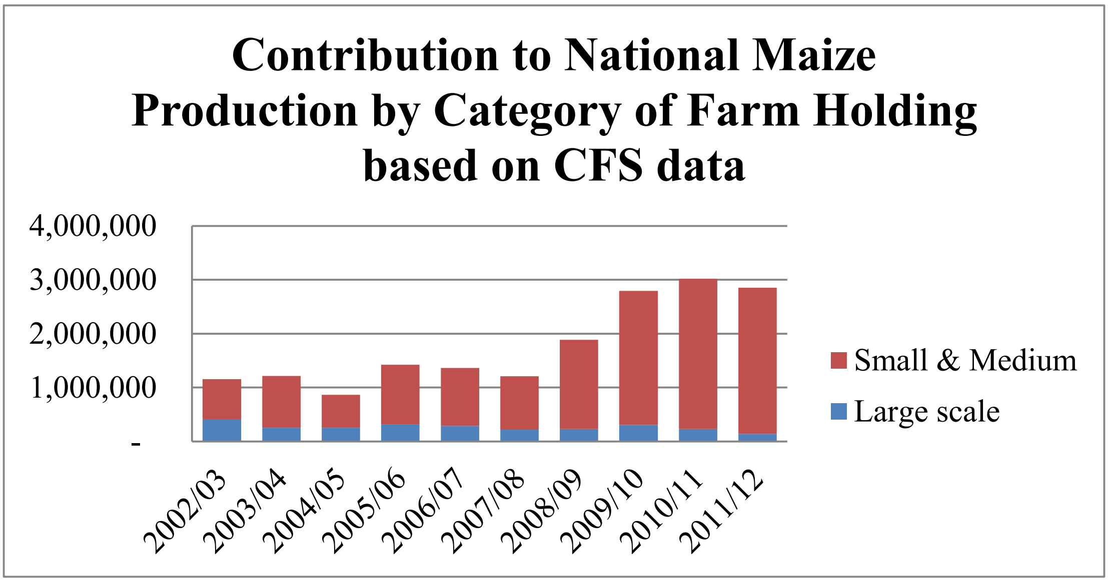
[3] 2008/2009 CFS Survey, CSO/MACO

[4] MACO analysis

patterns in Southern Africa, area planted to maize remains quite consistent.. In deficit years, imports usually meet the shortfall in national maize requirements.

Despite maize being widely grown by most Smallholder households, only about 25 percent of smallholder farmers in Zambia sold maize in both 1999/2000 and 2002/2003 seasons (Jayne et. al, 2007). Large scale farmers however, typically sell most of the maize they grow. The table below shows that the contribution to national maize production by large scale farmers has been trending downwards.

Figure 2: Contribution to National maize production by Large scale and small and medium scale farmers



Source: MACO/CSO Crop Forecast Survey 2003/04 to 2009/10

Most of the maize produced by large scale farmers is sold to milling companies who process it for the urban consumption market. Similarly over the past five years, the Government, through the Food Reserve Agency, has been increasingly purchasing more maize from smallholders. This maize is then sold to millers or in surplus years, exported to neighbouring countries.

Maize contributes significantly to the total calorie requirements of the average Zambian. According to the Zambia National Food Balance Sheet, in the 2007/2008 marketing season, maize was expected to provide 55% of total calorie requirements from staple foods. Cassava provided the second highest at 36%. However, the cassava figure is based on potentially available cassava throughout the season. Although the percentage is growing, very little cassava compared to maize is marketed nationally.

Table 1: Main Staple Contribution to Carbohydrate Requirements

Main Staple Contribution to carbohydrate requirements based on the 2007/2008 Zambia National Food Balance Sheet

Processed staple food for human consumption	Crop-to-food product conversion factor	Energy value (kCal/kg)	Required energy (kCal/capita/ day)	Required energy (%)
Mealie meal	0.9	3390	782	0.55
Rice	0.7	3350	16	0.01
Wheat flour	0.75	3400	77	0.05
Sorghum/millet fl.	0.9	3350	24	0.02
Other tubers	0.8	1100	14	0.01
Cassava flour	0.25	3200	507	0.36

Source: 2007/2008 FBS based on the 2006/2007 Crop Forecast Survey

The implication of this distribution is that maize is likely to have a much lower price elasticity of substitution compared to the other crops which constitute a much smaller share of the average Zambian's diet. A crop like rice is likely to have a relatively high price elasticity of substitution, based on the proportion it contributes to total energy requirements. The issue of elasticities is discussed in greater detail in the methods and analysis sections.

Problem Statement

There is widespread concern that higher food prices could severely jeopardize the food security situation in low-income countries where a large part of the population has very low purchasing power. According to IFPRI, the expansion of ethanol and other biofuels could reduce calorie intake by another 4-8% in Africa and 2-5% in Asia by 2020.[5] The "wage-good" nature of dominant staple foods in some countries could suggest broader economy-wide effects of higher food prices on wages and industrial competitiveness. However, the effects of a secular change in food prices depend ultimately on government policy responses, the structure of the internal economy, the ability of producers to respond to higher prices, and other local conditions. Maize prices like those of most commodities are generally exogenous to individual farmers.

It is not clear what impact global cereal price hikes will have on households given that smallholder agricultural households are non-homogeneous. Some households are able to increase production in response to changes in price whilst other households are unable to increase production on account of various constraints. This ability to increase production in response to price rises, also known as 'supply response' may be constrained by unavailability of additional land to expand cultivation, limited productive assets such as capital or labor, factor market failures, and agro-ecological conditions. For these reasons, structural change in food prices may advantage some farmers and disadvantage other rural households. The price increase might price some rural households out of the maize market. Rather than stimulating increased maize production, higher prices might increase net food insecurity. According to David Hallam, in a paper on agricultural supply in transition economies, knowledge of how agricultural supply is likely to respond to policy-induced price changes is self-evidently important in the

[5] The Economist 2008

definition and selection of appropriate price policies (Hallam). The main purposeof this thesis is to better understand the responsiveness of households to food price incentives and the various socio-economic factors that affect their responsiveness.

Heterogeneity within Zambian agriculture

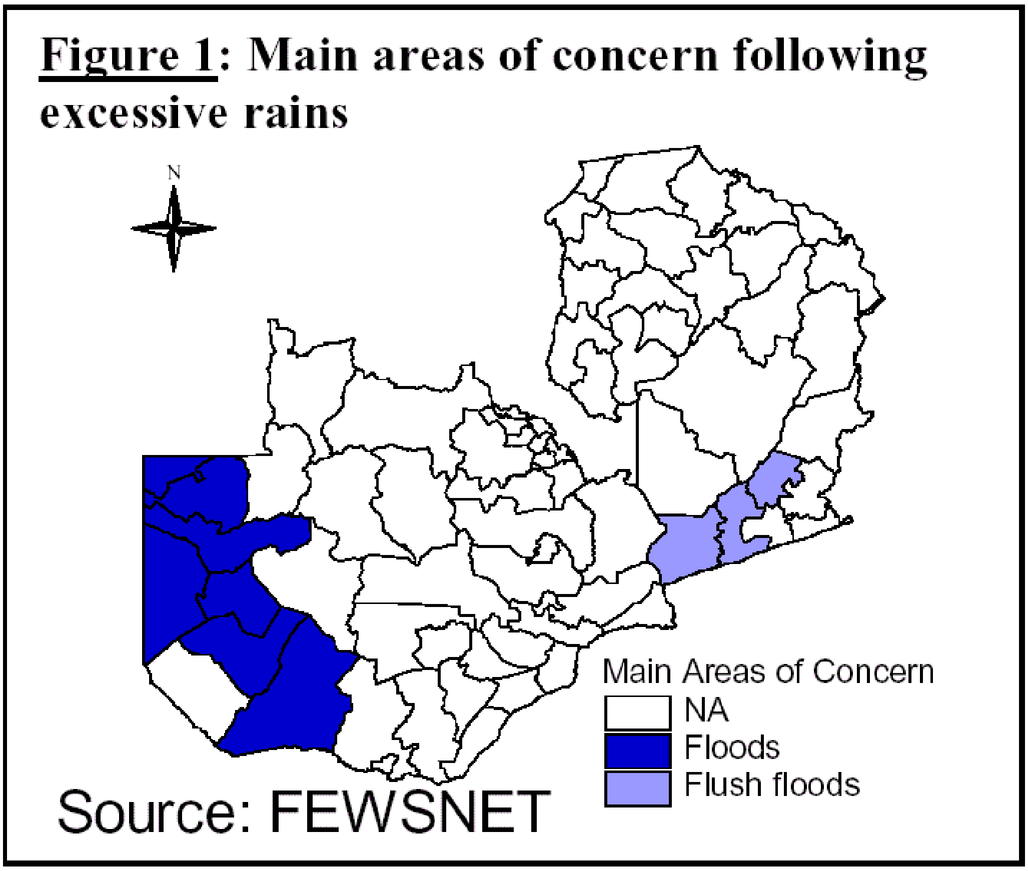
The Zambian agricultural sector is characterized by significant variation in the scope, size and capabilities of farm holdings. Land under cultivation by farm holdings ranges from 0.06 hectares to 16,000 hectares (CFS 2007/08)[6]. There are also considerable differences in the spatial characteristics of farm households and in the livelihood patterns rural households follow. 'Two percent of all smallholder farms nationwide accounted for over 40% of all the maize sold by smallholder households in Zambia in 2000/01 and 2003/04. This same two percent of smallholder households also accounted for about 17% and 20% of the total value of all crop sales of the smallholder sector (Jayne, 2007). Over the years (with the exception of the 2008/2009 marketing season) the price set by the FRA has tended to be considered by major buyers, as the de facto floor price for maize in a given season.

In some parts of the country, livestock rearing is more pronounced compared to other parts of Zambia. With the exception of maize, which is grown widely throughout the country, there some significant differences in the farming systems practiced around the country. Below is a map of Zambia showing some of the spatial disaggregation of livelihood patterns in Zambia.

[6] 2007/08 Crop Forecast Survey, conducted jointly between the Ministry of Agriculture and the Central Statistical Office

The blue shaded areas represent flood plains where livestock production is the main livelihood activity.

Figure 3: Map of Zones in Zambia



Source: Zambia Vulnerability Assessment Committee

To analyze the overall impact on Zambia's national food security, with specific emphasis on maize availability, requires measuring the distributional effects on the different categories of farmers in the country.

Table 2: National Estimate of Smallholders Growing Crops (Number of Households)[7]

Agricultural Year	Small-Scale	Medium-scale	% of hh that are Medium Scale	Total
2000/2001	760,983	22,259	2.9	783,242
2001/2002	765,323	25,566	3.3	790,889
2002/2003	1,002,298	24,788	2.5	1,027,086
2003/2004	946,672	43,169	4.6	989,841
2004/2005	1,127,418	44,030	3.9	1,171,448
2005/2006	1,148,470	40,386	3.5	1,188,856
2006/2007	1,126,386	48,349	4.3	1,174,735
2007/2008	1,101,219	44,610	4.1	1,145,829

Source: CSO/MACO/FSRP

The above data shows that at a macro level, there is some change in the number of households in each category. A goal of this thesis is to shed more light on the relationship between global maize price changes and the area planted to maize, given the set of diverse variables that confront households.

---

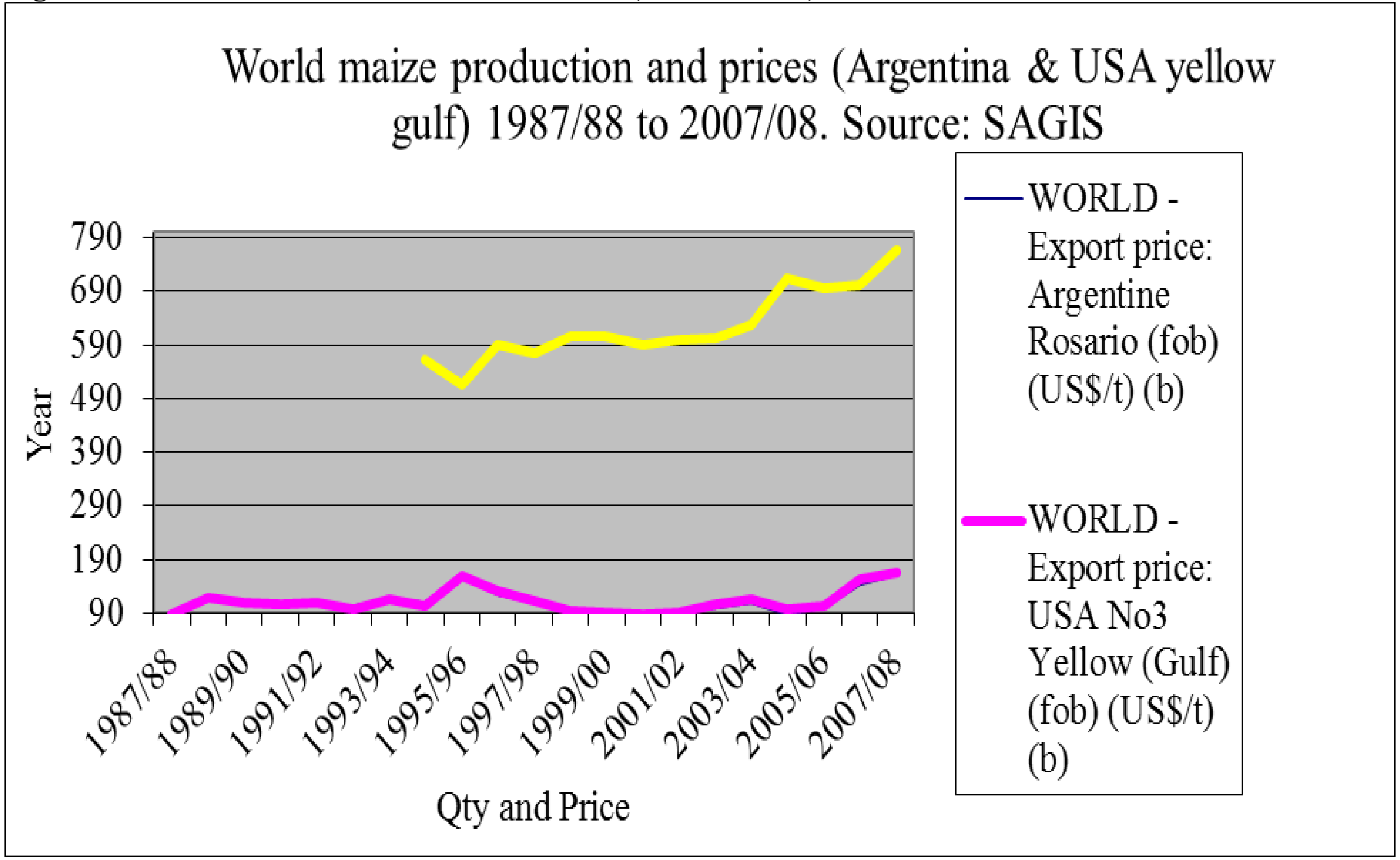
[7] The estimated national number of households growing crops is based on the weighting scheme used until 2007/2008. The observed jump in numbers from 2001/2002 to 2002/2003 is because of the introduction of the new weighting scheme based on the 2000 Census. Estimates prior to 2002/2003 are based on the 1990 Census weighting scheme.

Although global maize production has been increasing, prices have also been rising steadily especially over the last 3 years in part due to increased domestic demand for corn in the US, China and India. The year 2009 saw a relative stabilization of prices, in part due to the effects of the global economic crisis.

However, long term trends point towards a structural increase in food prices, in part due to population growth and the increasing prominence of the cultivation of bio-fuels as an energy source.

Consumption patterns in rural areas are also likely to change in response to higher maize prices. Recently, prices of oil based fertilizers have also trended upwards with the rise in crude oil prices. This has called into question the viability of fertilizer-dependent crops such as maize, especially among Smallholder farmers who have a limited asset base. It can be theorized that the compounded effect of input price shocks coupled with rises in maize prices may result in diversification of subsistence production to non-fertilizer dependent crops such as cassava which are also substitutes for maize.

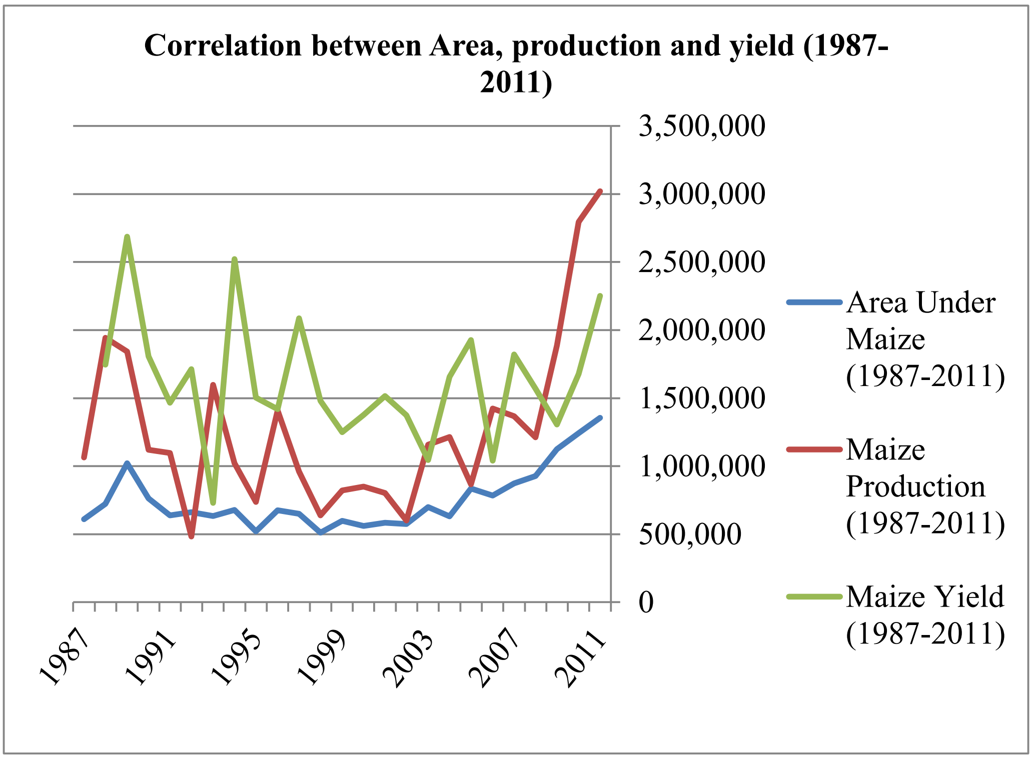
Figure 4: World Maize Production and Prices (1988 to 2008)



Source: South African Grain Information Service

Historically, a major challenge in improving farmer supply response has been the very low yields obtained by most small and medium scale farmers in Zambia. Based on the crop forecast survey data, yields over the past 20 years have averaged 1.53 mt/ha. In addition, there is more correlation between maize production and yields, compared to area planted and yields suggesting that factors other than the size of area that farmers choose to plant do explain the variation in production. One such factor is the rainfall patterns. It is therefore important to also control for factors that are outside farmers' control such as the weather when estimating the degree of farmers' responsiveness to prices. This is done by creating both yield model and an area planted model. Details are discussed in the methods section.

Figure 5: Crop Forecast Survey Area, Production and Yield for Maize 1987 - 2008



Thesis Objectives

This thesis aims to conduct a detailed analysis of the production response to a rise in maize prices in Zambia among rural households. The smallholder sector is analyzed. The primary objective of the study is to measure how maize production will respond to structural changes in price with a focus on the distributional effects of rising prices of maize, given the great heterogeneity within the smallholder sector as well as rising costs of production. Specifically, the study's objectives are as follows:

1. To estimate the supply response of the different categories of farmers namely small and medium scale (growing between 0 – 19.99 hectares of crops);
2. To conduct a simulation to examine the differential impact of rising maize prices on small holder producers. What have the different categories of producers done in response to changes in maize prices over the period under analysis?

3. To compute elasticities to measure the relationship between maize prices and other variables of the production function on area and yield.

Hypotheses

H1: Not all small holders will benefit from a rise in maize prices.

Conventional wisdom suggests that farmers prefer higher commodity prices to lower prices. Since maize is the most widely grown crop in Zambia, it is assumed that increases in the price of maize will have widespread positive benefits to growers.

H2: Elasticity of supply will be positively related to landholding and asset size.

Farmers with more access to land, ploughs, and draught power are better able to take advantage of higher prices of maize and have a faster production response.

H3: In the short run, higher input prices (especially fertilizer) have a higher impact on maize production compared to a proportional increase in maize prices.

Before a production decision is made, farmers take into account the cost of inputs, which they generally know upfront, and some expectation of the price they will receive for their crop. What is the substitution effect of maize production with crops such as cotton and cassava?

Literature Review

Previous studies have used several approaches to measure price policy changes and the resulting impacts on households or the agricultural sector in general. Each has advantages and disadvantages. The Organization of Economic Cooperation and Development (OECD) conducted a series of case studies to analyze the effect of agricultural and trade reform. According to the OECD, these studies produced two main findings specifically in developing countries, The first was that market interventions often produce ambiguous effects on the distribution of income, and in poor countries it typically is impossible to use a price intervention to make some poor households better off without making other poor households worse off (Brooks J, et al 2008) This trade off results from the joint estimation of the impact of price changes on both supply response and consumption decisions.

A case study approach was also used in the US, where the impact of rising grain prices was already being felt by many farm enterprises. Some farmers who raise livestock and grow corn were reportedly increasing the production of corn at the expense of livestock and other crops (Meating place, 2007). The structure of US (and other developed countries) agricultural systems differs considerably from those in developing countries in several key areas. Farmers in developed countries are generally commercialized compared to most farmers in developing countries such as Zambia. They do not rely on own harvest for consumption requirements. Consequently, price changes may have a different impact on supply response and this difference manifests itself through the profit effect, which is discussed in more detail in the conceptual framework section.

A case study approach would be more suited to the large scale farming sector in Zambia, since the large scale sector is non-subsistent and has a structure that more closely mirrors western farm organization. In addition, the large scale sector contributes a relatively smaller share of total national maize output. A more thorough treatment of the topic of supply response has to look at the small-holder sector and take account the organizational models that obtain in small holder agriculture, particularly the dichotomy of production and consumption decision making by subsistence households.

In a 2008 Policy Research Working Paper, Martin and Ivanic use Living Conditions Survey Data from 10 countries to measure the net effect of higher prices, raising the real income of those selling food but at the same time hurting food consumers many of whom are really poor (Martin & Ivanic, 2008). Martin and Ivanic measured short run impacts on household income and costs of living following the changes in food prices. They use household survey data which has both consumption and production information of the main food commodities. They then measure the change in household real income and also estimate the impact of food prices on poverty rates and poverty gaps. In the first experiment, a simulation measuring the effect of a 10 percent change in prices is conducted. This simulation uses cross sectional data analysis. An assumption of international price transmission to domestic prices is made. Martin and Ivanic also measured the impact of commodity price changes on changes in the wage rate for unskilled labour.

This thesis goes beyond the work of Martin and Ivanic by focusing more on the differential impacts of price changes on heterogeneous rural households' supply response. Consumption and demand dynamics are excluded from the analysis in this thesis. The Martin and Ivanic study takes a short run, cross-sectional approach to measuring supply response. Structural changes, by definition typically have long run consequences on livelihood patterns. This thesis looks at household supply response over a period of 8 years, from 2001 to 2008. It is hoped that the long run supply response will provide more insight into households long-term ability to cope with price changes. This analysis also uses panel data to control more effectively for unobserved time-constant factors correlated with prices.

Supply Response and Production/Consumption Decision making by subsistence Households

Some agricultural households are expected to raise production in response to higher prices of maize. However, households who also rely on off-farm labour to meet part of their consumption requirements may have to put off investment in increased future production in favour of off-farm labour activities to meet their current consumption requirements. Consequently variations in the price of major crops will frequently affect both producers and consumers (Squire, 1980). At question in this thesis, is the net effect on production.

In The Analysis of Household Surveys, Angus Deaton provides two examples that measure the effects of change in rice prices on the distribution of income in Thailand, using the social economic survey of 1981-82 (Deaton, 1997) as well as the impact of a social pension on reaching poor households in South Africa. The South African Living Standards Survey was used for the analysis and Deaton focuses on a cross-sectional analysis (and consequently short-run measurement) of the effect of price changes on household welfare. This thesis differs from the work of Deaton and Strauss, et al by focusing exclusively on long run effects on supply.

In Zambia, research results indicate 'that about 40-45% of the total marketed supply of maize from the smallholder farm sector was produced by only 2 percent of the smallholder farms (Jayne et al), indicating a very high concentration of the marketed surplus. Are household maize sales correlated with income and wealth? More farm households are buyers or net buyers of maize than sellers implying that the majority of small-scale farm households may be adversely affected by price and trade policies designed to raise market prices of maize'[i] It is important to stress the need for a disaggregated estimation of the effects of maize price changes that differentiates the supply response across heterogeneous types of small holders.

The analysis in this thesis focuses exclusively on supply response. Demand systems are not incorporated in the analysis. Consequently, the results paint a partial picture of the food security situation in Zambia as a result of increasing maize prices. The non-inclusion of demand systems allows for a more robust treatment of supply response among the broad range of variables included in the models.

Existing data suggests that the supply response of households to changes in the price of cash crops can be influenced by the gender of the household. According to Whitehead decision making in households is not necessarily "joint," and individual control over resources is valued by household members; and preference heterogeneity between spouses can have real consequences for changes in households' production, income, and welfare accompanying change in their economic environments (Whitehead, 1990).

Benefits from changes in prices might not be distributed proportionately to household members. This may in turn influence future cropping decisions. The primary models used in this analysis incorporate the gender variable interacted with the maize price. This enables an understanding of how the supply response to price changes will vary between male and female headed households.

Data

In order to estimate a supply response model, it is necessary to have a set of data on variables related to the factors of production; land holding size, household and hired labour supply (possibly broken down by sex), farm and non-farm inputs, purchased and household supplied variable inputs, fixed farm assets, basic demographic characteristics, and prices for production inputs, including wages (Strauss, 1980). The supplemental survey datasets do contain most of the required variables.

The data used in this thesis is from three major sources: (1) nationally representative longitudinal panel supplemental surveys (SS) to the 2000 Post harvest survey (PHS); (2) annual post-harvest survey data; (3) qualitative data obtained from interviews with selected key informants from the agricultural sector in Zambia (4) Weather data from the Zambia Meteorological Department.

Household level variables

The Ministry of Agriculture and Co-operatives (MACO) in collaboration with the CSO, conducts annual Crop forecast and Post-harvest surveys in all districts of Zambia. Separate samples are drawn for small and medium scale holdings (commonly referred to as Smallholders) and for large-scale holdings. A complete enumeration is conducted of all large-scale holdings in Zambia. However, due to the lack of a reliable large scale frame as well as operational challenges, the omission of a few large scale holdings from the survey does occur.

In 2000, the annual Post Harvest Survey was conducted jointly by MACO/CSO with a sample size of 20 households interviewed in each of the 390 selected Standard Enumeration Areas in the country (CSO Training Manual, 2007).[8] The 2000 PHS collected information on the 1999/2000 farming season. However, the main focus of the PHS has typically been on the collection of information related to production of agricultural commodities. Little emphasis is given to collection of data on non-farm income sources, which may have an impact on production related decisions. This is part of the motivation behind the supplemental survey. The survey was designed to collect data that would supplement information collected in the PHS. The first supplemental survey was conducted in 2001 and revisited the same households interviewed in the 1999/2000 PHS. The reference period for most of the information captured in the SS was still based on the 1999/2000

[8] A standard Enumeration Area (SEA) is the smallest sampling cluster used for sampling by the CSO. SEAs are part of Census Supervisory Areas (CSAs) which typically have 4-6 SEAs each. All stage one sampling of two stage sampling designs used in Zambia are based on SEAs. A detailed explanation of the sample selection is given in the appendix.

farming season. In total three SS have been conducted to re-interview the same households first interviewed in 1999/2000. The three supplemental surveys were longitudinal surveys conducted for smallholder households and the same households were visited in 2001, 2004 and 2008. The resulting dataset is a panel dataset at household level. This allows comparisons to be made at household level over the eight year period. [9].

The CFS, PHS and SS surveys use the national census conducted every 10 years to derive the sampling frame from which the 390 SEAs are selected. Because the SS is a longitudinal panel survey, it is based on the 1990 census. A sample of size 7,880 small-scale households is drawn. 'About 96% of the farms in these nationally representative surveys are in the small-scale (0.1 to 5.0 hectare) category, with the mean area per small-scale farm being 1.4 hectares. About 4% of the farms are in the "medium-scale" category. For ease of citation, we refer to the full sample of both categories as the "Smallholder" farming sector. (FSRP)

---

[9] The 2001 supplemental survey did not collect any information on the fields under management by the household during the reference period. This was collected during the preceding PHS. Since the two surveys shared the same households and reference period, field level data from the PHS was incorporated into the SS. The same households selected for interview during the PHS in 1999/2000, were re-visited during the 2001 supplemental survey. The 2001 supplemental survey did not collect any information on the fields under management by the household during the reference period. This was collected during the preceding PHS. Since the two surveys shared the same households and reference period, field level data from the PHS was incorporated into the SS

Sample Size and Attrition

Of the 6,922 households interviewed in 2001, 5,420 were re-interviewed in May 2004 (Chapoto and Mason, 2007). The number of households re-interviewed in the 2008 survey dropped to 4570 households.

In order to maintain the sample size after taking attrition into account, the 2008 survey modified the sample design to enable replacement of non-contact panel households.[10] However, since this analysis is conducted on the balanced panel interviewed in each of the three years, we do not include replacement households in the analysis. Yamano and Jayne (2005) propose a re-interview model that can be written as

$P(R_{it} = 1) = f(HC_t, ET_{it}, PD)$ where R_{it} is equal to 1 if a household I was re-interviewed at time t, conditional on the household being interviewed in the previous period, and zero otherwise; HC_t is a set of household characteristics in the initial survey, ET_{it} is a set of enumeration team dummies, which in this thesis have been replaced with provincial dummies.

Table 3: Attrition rate for household interviews in Supplemental Surveys

Supplemental Survey Year	Number of Households interviewed	Re-interview rate (Percentage)
2001	6,922	na
2004	5,420	78
2008	4,570	66

[10] If less than 20 of the original sampled panel households in the SEA were captured during listing, replacement households were selected to be interviewed, to bring the total number interviewed in each SEA to 20.

Source: Supplemental Surveys, 2001, 2004 & 2008

Two probit attrition models with probability of the household being re-interviewed in 2004 and 2008 are included. Household Characteristics include province dummies, soil-type in that area, quartiles of total land under the control of the household in, total household size, sex of household head, total cattle raised by the household, asset index of 7 assets recorded in the survey, a dummy variable indicating whether the household head was related to the village head at time of land allocation, the number of household members in formal employment, soil-type interacted with time, age of household head, age of spouse, education level of head, education level of spouse, education level of highest educated other member and dependency ratio[11].

Total household size, number of members in formal employment and the eastern province dummy produced statistically significant estimates for the systematic difference between households that were interviewed in 2004 versus those that were not interviewed.

In the 2008 survey, household size, household asset index, number of members in formal employment, education level of household head, as well as the eastern and Luapula provincial dummies produced statistically significant estimates for the systematic difference between households that were interviewed in 2008 versus those that were not interviewed. Inverse Probability Weighting is proposed to correct for bias, in cases where it is

[11] Dependency ratio is typically computed as number of household members younger than 15 and older than 60 divided by the number of household members between 15 and 60. However, there are several households in the dataset with no members between 15 and 60. To overcome this mathematical constraint, total household size is used as the denominator.

identified as a problem. The main model in this thesis use the *"xtivreg"* command. This model offers correction for potential bias problems. Other work by Mather, Boughton and Jayne suggests that attrition bias is not a problem in this panel dataset. Consequently, correction for attrition bias is not done.

Table 4: Attrition bias results test

Zambia (using 2000 data)	p-value for test of H0: β*re-interviewi,t + 1 = 0 vs. H1: β*re-interviewi,t + 1= 1	
Auxiliary regressions		
Quantity of maize sold (kg)	Tobit	0.17
Farm-gate maize sale price (LC/kg)	OLS	0.228
Quantity of subsidized fertilizer received (kg)	Tobit	0.24
Quantity of fertilizer used on maize (kg/ha)	Tobit	0.2
1=HH used improved variety	Probit	0.006
Maize market participation regressions		
1=HH sold maize	Probit	0.018
ln(Quantity of maize sold (kg))	Log Normal	0.046

Mather, Boughton, Jayne,

Table 5: Probit model to measure attrition against household characteristics in 2004 & 2008

VARIABLES	Was hh re-interviewed in 2004?	Was hh re-interviewed in 2008?
1st quartile of total area of land held by hh interacted with time	0.0958	0.000605
	(0.102)	(0.0960)
3rd quartile of total area of land held by hh interacted with time	0.202**	0.185**
	(0.0960)	(0.0896)
4th quartile of total area of land held by hh interacted with time	0.187**	0.126
	(0.0932)	(0.0867)
5th quartile of total area of land held by hh interacted with time	0.142	0.124
	(0.0921)	(0.0862)
total household size	0.0331***	0.0222***
	(0.00962)	(0.00861)
sex of hh head	0.00957	-0.00716
	(0.180)	(0.165)
total cattle raised by hh	0.00189	-0.00340
	(0.00382)	(0.00260)
asset index of 7 assets listed in survey	-0.00208	-0.00438***
	(0.00134)	(0.00121)
hh related to village head at time of land allocation	0.114*	0.0900
	(0.0615)	(0.0560)
number of hh members in formal employment	-0.468***	-0.531***
	(0.111)	(0.106)
soiltype interacted with time	0.00685	0.00145
	(0.00518)	(0.00472)
age of household head	0.00260	0.000518
	(0.00371)	(0.00343)
age of spouse	0.00246	0.00275
	(0.00419)	(0.00388)
education level of head	-0.0114	-0.0268***
	(0.00938)	(0.00865)
education level of spouse	0.00639	0.00584
	(0.0104)	(0.00969)
education level of highest educated other member	0.0149	0.00761
	(0.0117)	(0.0109)

Table 5 (cont'd)

dependency ratio-hh members between 15 and 60 over hh size	0.0455	-0.0774
	(0.171)	(0.158)
central province	0.0798	0.0772
	(0.159)	(0.148)
copperbelt province	-0.00132	0.137
	(0.173)	(0.164)
eastern province	0.322**	0.298**
Table 5 continued	(0.158)	(0.146)
luapula province	0.163	-0.377**
	(0.164)	(0.151)
northern province	-0.0129	0.0507
	(0.155)	(0.146)
north western province	-0.194	-0.125
	(0.173)	(0.165)
southern province	-0.00245	0.149
	(0.156)	(0.147)
western province	-0.0463	-0.0900
	(0.165)	(0.154)
Constant	0.146	0.278
	(0.288)	(0.265)
Observations	2990	2990
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1		

Two groups of households were generated with households cultivating less than 5 hectares classified as category A. Those cultivating 5 hectares or more were classified as category B households. Ten households were to be selected from each category to generate the required sample of 20 households per SEA. Large scale PHS data is obtained by complete enumeration. This data is then aggregated to district level and merged with the district level survey data after the district level small and medium scale data has been boosted with the appropriate weights file. Large scale data was aggregated to national level.

Price Data

Naïve price expectation theory assumes that a household makes planting decisions based on prices it received in the previous season. The supply response model used in this thesis uses naïve price expectation theory. The supply response models include staple commodity prices as explanatory variables. The supplemental surveys did collect household level price data on sales of commodities. Household price data provides a much richer pattern of responses, due in part to the relatively high number of responses and the household level variation in the data. However, there are drawbacks to using household price data. Deaton notes that unit values (quantity purchased divided by expenditure on that unit) are affected by choice of quality as well as by the actual prices that consumers and producers face in the market (Deaton 2006). Measurement error is another problem that arises from household level price data. However these prices were post-harvest and are not known at planting time. Consequently, the supply response model cannot use the supplemental survey household price data.

The annual PHS survey does collect information on prices of crops sold by the household. District level prices are obtained by aggregating household price data. The analysis then picks the aggregated PHS district level price data for the season prior to when planting was actually done.

Table 6: Nominal Median Provincial Producer Maize Prices K/kg)

Year	Central	Cbelt	Eastern	Luapula	Lusaka	Northern	Nwestern	Southern	Western
1999	447.76	470.59	347.83	406.96	398.01	417.39	398.01	417.39	434.78
2002	521.74	591.30	434.78	500.00	521.74	521.74	521.74	470.59	608.70
2006	521.74	608.70	521.74	588.24	643.48	641.85	626.09	521.74	634.78

Source: MACO/CSO PHS data

Table 7: Deflated Median Provincial Producer Maize Prices (K/kg)

Year	Central	Cbelt	Eastern	Luapula	Lusaka	Northern	Nwestern	Southern	Western
1999	100.00	105.10	77.68	90.89	88.89	93.22	88.89	93.22	97.10
2002	116.52	132.06	97.10	111.67	116.52	116.52	116.52	105.10	135.94
2006	116.52	135.94	116.52	131.37	143.71	143.35	139.83	116.52	141.77

Source: MACO/CSO PHS data

Using aggregated district price data does result in some loss of household price variability, which might be important in understanding household decision making dynamics. However, because the household-specific prices in the Supplemental Surveys are post-harvest prices and not known at planting time, it is necessary to resort to more aggregated district-level prices (available in the Post-Harvest Surveys) from the prior years under the assumption of naïve expectations. The median maize price from households selling maize in the prior season is used to obtain aggregated district-level prices.

Fertilizer Price Data

Fertilizer data from the supplemental surveys was used in the model analysis. Households reported quantities of fertilizer used as well as price paid for fertilizer during the largest acquisition transaction. The household level prices reported for the largest transaction were aggregated to national level.

The annual national household reported price shows a decline in the real price over the period. Part of the explanation for this could be the effect of international price transmission. Fertilizer prices are based on oil prices and tend to fluctuate accordingly.

Adult equivalents were calculated and used in the model in place of the household size variable.

Measurement of Asset Variable using Principal Components Analysis

The 2001, 2002 & 2008 surveys collected information about type, quantity and values of asset holdings by households . Two main data challenges are noted with the asset data

i. The 2001 survey collected values only for three asset types (ploughs, harrows and ox-carts only). The two subsequent surveys did collect a more comprehensive set of asset variables. However, panel analysis asset values across the eight year period is rendered difficult given the lack of values for assets outside the set of three mentioned above. In order to overcome this constraint, binary variables (YES/NO responses for whether the household owned the asset) relating to ploughs, harrows, oxcarts, pumps, trucks, cars, bikes, bicycles and mills are introduced.

ii. According to Filmer and Pritchett ranking households based on economic status measures such as asset holdings, requires a normalizing or weighting procedure to eliminate bias (Filmer & Pritchett, 1998). Among several methods for identifying the appropriate weighting scheme, Principal Components Analysis is proposed as the most effective in relation to the data structure.

  Principal Components Analysis is a technique for extracting from a set of variables those few orthogonal linear combinations of the variables that capture the common information most successfully (Langyintuo, 2008). The assets being included in the analysis need to be recorded as binary indicators only, with no values included in the analysis. In the case of the asset holdings in the supplemental survey datasets, the use of PCA is pragmatic response to a data constraint problem[ii]. Filmer and Pritchett further argue that the resulting asset index generated using PCA must be viewed as a proxy for a households long-run economic status.

  PCA starts by specifying each variable normalized (weighted) by its mean and standard deviation. For example,

$$a_{1j} = (a_{1j}^{*} - a_{1}^{*}) / s_{1}^{*},$$

where $\overset{*}{a_1}$ is the mean of $\overset{*}{a_{1j}}$ and $\overset{*}{s_1}$ is its standard deviation.
The selected variables are expressed as a linear combinations of a set of underlying components for each household j:

$$a_{1j} = v_{11}A_{1j} + v_{12}A_{2j} + \dots v_{1K}A_{Kj}$$
$$\dots \qquad \forall_j = 1,\dots, j \qquad (1)$$
$$a_{Kj} = v_{K1}A_{1j} + v_{K2}A_{2j} + \dots v_{KK}A_{Kj}$$

Where the *As* are the components and the *v*s the co-efficient on each component for each variable (and do not vary across households). PCA finds the linear combination of the variables with maximum variance usually the first principal component A_{1j} and then a second linear combination of the variables orthogonal to the first, with maximal remaining variance, and so on.

The 'scoring factors' from the model are recovered by inverting the system implied by equation (1), and yield a set of estimates for each of the K principal components:

$$A_{1j} = f_{11}A_{1j} + f_{12}A_{2j} + \dots f_{1K}A_{Kj}$$
$$\dots \qquad \forall_j = 1,\dots, j \qquad (2)$$
$$A_{Kj} = f_{K1}A_{1j} + f_{K2}A_{2j} + \dots f_{KK}A_{Kj}$$

The first principal component, expressed in terms of the original (un-normalized) variables, is therefore an index for each household based on the expression:

$$A_{1j} = f_{11}(\overset{*}{a_{1j}} - \overset{*}{a_1})/(\overset{*}{s_1}) + \dots + f_{1k}(\overset{*}{a_{kj}} - \overset{*}{a_k})/(\overset{*}{s_k})$$

(3)

The assigned weights are then used to construct an overall 'wealth index', applying the following formula:

$$W_j = \sum_{i=1}^{k} [b_i(a_{ji} - x_i)] / s_i$$

(4)

Where: W_j is a standardized wealth index for each household; b_i represents the weights (scores) assigned to the (k) variables on the first principal component; a_{ji} is the value of each household of the k variables x_i is the mean of each of the k variables; and s_i is the standard deviations.[12]

The interpretation of results of the above manipulation is that a negative index, implies that the household is poorly endowed relative to the community whilst a positive W_j means the household is relative well off.

Included in the thesis analysis are binary variables for ploughs harrows oxcarts pumps trucks cars bikes bicycles and mills. An critical assumption is that having the asset is important in influencing production decisions, the quantity held is not. An additional output from the creation of the household level wealth index is the calculation of the impact factor. This is obtained by dividing the individual asset score by the corresponding standard deviation.

---

[12] Extracted from Langyintuo. For a thorough treatment see both Langyintuo (2008) and Filmer & Pritchett (2001).

Table 8: Impact Factor, 2001, 2004 & 2008 Supplemental Survey

	Plough	Harrow	Oxcart	Pump	Truck	Car	bike	bicycle	Mill
Mean	0.186	0.047	0.0913	0.0101	0.0063	0.0112	0.008	0.5088	0.0145
Standard Deviation	0.3891	0.2116	0.28881	0.0999	0.0792	0.105	0.0892	0.4999	0.1197
Score	0.7505	0.708	0.7684	0.4155	0.3618	0.4049	0.0341	0.3315	0.3826
Impact Factor	1.93	3.35	2.67	4.16	4.57	3.85	0.38	0.66	3.2

An impact factor of 1.93 for a plough indicates that the household's relative wealth ranking will adjust by 1.93 if the household acquires a plough. The significance of the impact factor is in comparing wealth index adjustments across the different types of indicated assets.

Access to Extension Services

Extension services are an important variable in influencing household productivity. All three surveys did collect household level information on what type of extension advice the household used. However, including the type of advice used by the household as an explanatory variable would lead to endogeneity. A variable asking about the availability of extension services is more appropriate. To proxy availability, the use of extension advice by the household was converted to a binary variable and aggregated to SEA level. The aggregated variables were then converted to a percentage of households accessing advice. This acts as a proxy for the availability of advice in an area. The three types of advice considered are minimum tillage, crop rotation and use of crop residues.

Water Requirement Satisfaction index

Consistent and official rainfall data in Zambia is available for a limited number of rainfall stations (approximately 40) out of the 72 districts (Zambia Meteorological Department). Obtaining district level estimates of rainfall becomes a challenge, especially for districts without any official rainfall estimates reporting system. Satellite estimates of rainfall are however, available for Zambia.

The meteorological service in Zambia has a process of interpolation to 'ground truth' the rainfall estimates from the satellite imagery with the actual rainfall figures on the ground. The resulting estimates are used in several of the crop performance analysis tools being used to forecast crop production. The process of interpolation involves comparing the satellite estimate of rainfall in a particular location with the actual quantity collected from the weather stations in that district. The difference between the satellite estimate and the actual estimate on the ground is then extrapolated to adjust the satellite data for areas without ground estimates. One such use of interpolated data is the Water Requirement Satisfaction Index. This is an indicator of crop performance based on the availability of water to the crop during a growing season. According to the US Geological Survey, WRSI for a season is based on the water supply and demand a crop experiences during a growing season. It is calculated as the ratio of seasonal actual evapo-transpiration (AET) to the seasonal crop water requirement (WR):

$$WRSI = (AET / WR) * 100. \qquad (1\text{-}10)$$ [13]

Actual evapotranspiration or AE is the quantity of water that is actually removed from a surface due to the processes of evaporation and transpiration. Crop Water Requirement is calculated based on the type of crop grown and the stage of crop growth.

WR is calculated from the Penman-Monteith potential evapo-transpiration (PET) using the crop coefficient (Kc) to adjust for the growth stage of the crop:

$$WR = PET * KC.$$

The spatially explicit water requirement satisfaction index (WRSI) is an indicator of crop performance based on the availability of water to the crop during a growing season. The WRSI data is used in the analysis.

[13] The WRSI is expressed as a percentage

Methodology

Conceptual Econometric Approach

This thesis focuses exclusively on the supply response of households to changing maize prices. However, in a true subsistence household, production and consumption decisions will be made jointly. The household must, to the extent possible, produce what it intends to consume. The household relies on its own labour and asset endowment to meet its production requirements.

Separability versus Non Separability

Singh, Squire and Strauss reviewed several different models that analyzed the dynamics of rural agricultural households. Their basic framework suggests that 'a large part of agriculture, however, is made up of semi-commercial farms in which some inputs are purchased and some outputs are sold. In these circumstances, producer, consumer and labor supply decisions are no longer made simultaneously, although they are obviously connected because the market value of consumption cannot exceed the market value of production less the market value of inputs.'(Singh, 1986) If a household is a price taker, then the production decisions it makes are likely to try and maximize production since the household cannot influence the price by changing its output. Non-agricultural households maximize their utility by maximizing output and using the resulting income to purchase their consumption requirements. Agricultural households on the other hand may devote some of their land to producing crops meant for consumption. This decision might not necessarily be motivated by the anticipated price of maize. In short, there is often a relationship between production and consumption decisions made by rural agricultural households especially more subsistence households.

In order to assume separability however, Deaton suggests that an assumption of perfect labour markets must be made. This means that there must be no difference between the household working on its own farm and the household selling its labour in the labour market. Household labour and hired labour must be perfect substitutes. By definition recursive models are hierarchical i.e. all causal effects in the model are unidirectional in nature (Williams). This means that the first endogenous variable is affected only by the endogenous variables. In a household model, specifically in relation to the production and consumption decision making process, this assumption can break down due to a net increase in the profit effect; an increase in the price of maize will, *ceteris paribus*, increase the household's profit. The increased profit can be used to increase the household's consumption of maize. On the other hand, the household can decide to maximize short term returns at the expense of medium term production decisions and expend more effort in off farm labour in order to meet its short term food requirements. In such circumstances, the net effect of a price rise is to reduce future production.

In order to address the complications arising from the use of joint estimation models which are primarily caused by data gaps in consumption data collected by the supplemental survey, the analysis is simplified and restricts itself to supply response changes. Consumption decision making and the profit effect are however discussed in the supply response model development in order to provide some context for the results discussion section. The lack of inclusion of joint production/consumption decision making in the main model is due to insufficient data availability from the three surveys.

The broader question of the net impact of maize price changes on the nation as a whole requires several other important concepts to be addressed. The issue of rural versus urban dichotomy also presents considerable analytical challenges. In addition, the structural separation of large scale production dynamics versus small holder dynamics adds further challenges. Restricting the analysis to supply response helps to break down this research topic into more meaningful policy considerations that are more easily comprehensible to the policy makers.

Other important conceptual considerations include;

- Government policy and its impact as a signal
- Price expectation theory
- Price substitution
- Endogeneity

Government policy is incorporated in the model by use of some variables. Since the price at which the farmers will sell their commodity is not known at the time of planting. The previous year's price is used as an explanatory variable. Prices of substitute crops are included in the model. Endogeneity is addressed through the use of instrumental variables.

Micro-Economic theory of the effects of price changes on producers

The agricultural model used in this thesis is based partly on the theory presented by Singh, Squire and Strauss. A general model is discussed and justification for the type of model used in the analysis is given. This is based on some basic empirical analysis of the data.

A Basic model of agricultural household Behaviour

For an agricultural household producing one crop, facing exogenous prices for inputs and outputs, and using one variable input, a typical utility (U) function can be represented as;

$$U = (X_a, X_m, X_l) \qquad (1\text{-}1)$$

Where the commodities are an agricultural staple, (X_a), a market purchased good (X_m), and leisure (X_l). The maximization of utility is subject to a cash income constraint:

$$p_m X_m = p_a(Q - X_a) - w(L - F)$$

Where p_m and p_a are the prices of the market purchased commodity and the staple, respectively, Q is the household's production of the staple (so that $Q - X_a$ is its market surplus), w is the market wage, L is the total labour input, and F is the family labour input (so that L-F, if positive, is hired labour and, if negative, off farm labour supply).

Other constraints faced by the household include;

X_l+F = T, a time constraint where T is the total stock of household time,

Q =Q(L, A)

A production constraint or production technology that depicts the relation between inputs and output, A is the households fixed quantity of land. Collapsing the three constraints into a single constraint and substituting the time constraint into the cash income constraint for Q and substituting the time constraint into the cash income constraint for F yields a single constraint of the form

$$p_m X_m + p_a X_a + wX_l = wT + \pi \qquad (1\text{-}2)$$

Where $\pi = p_a Q(L, A) - wL$ and is a measure of farm profits.

The Left hand side (LHS) shows total household 'expenditure' on market purchased commodity, households own output and time. The right hand side (RHS) is the value of the stock of time owned by the household.

Equations (1-1) and (1-2) are central to the analysis of agricultural household behaviour.

One measure used for farm profits is;

$$\pi = (p_a Q - wL)$$

In order to maximize each of these choice variables, we use the first order conditions for each choice variable.

For example, the first order condition for the labour input is;

$$p_a \delta Q / \delta L = w \qquad (1\text{-}3)$$

A household will equate its marginal revenue product of labour to the market wage. This equation contains only one endogenous variable, L. the other endogenouse variables do not influence the household's choice of L. Equation (1-3) can be solved for L as a function of prices (p_a and w), the technological parameters of the production function, and the fixed area of land.

The implication of this result is that production decisions can be made independently of both consumption and labour supply (or leisure) decisions. It further means that the model specification must exclude endogenous variables apart from the labour. All prices used must be considered exogenous.

The solution for L is

$$L^{*} = L^{*}(w, p_{a}, A) \qquad (1\text{-}4)$$

This solution can then be substituted into equation (1-2) to give

$$p_{m}X_{m} + p_{a}X_{a} + wX_{l} = Y^{*}$$

Maximizing utility subject to this new version of the constraint yields the following first-order conditions:

$$\begin{aligned} &\delta U / \delta X_{m} = \lambda p_{m} \\ (1\text{-}5) \qquad &\delta U / \delta X_{a} = \lambda p_{a} \\ &\delta U / \delta X_{w} = \lambda p_{w} \end{aligned}$$

And $p_{m}X_{m} + p_{a}X_{a} + wX_{l} = Y^{*}$

These are standard conditions from consumer-demand theory.

The solution to equation 1-5 yields standard demand curves of the form

$$X_{i} = X_{i}(p_{m}, p_{a}, w, Y^{*}) \qquad (1\text{-}6) \quad i = m, a, l$$

Demand depends on prices and income. For agricultural households, income is determined by the household's production activities. Changes in factors influencing production will change Y^{*} and hence consumption behaviour. This is the recursive property of the model. However, the level of 'recursiveness' ultimately is an empirical question that depends on the characteristics of households.

What happens when the price of the agricultural staple is increased? From equation (1-6)

$$\frac{dX_{a}}{dp_{a}} = \frac{\delta X_{a}}{\delta p_{a}} + \frac{\delta X_{a}}{\delta Y_{a}}\frac{\delta Y^{*}}{\delta p_{a}} \qquad (1\text{-}7)$$

The first term on the RHS is the standard result of consumer demand theory. The second term captures the profit effect for a normal good. A change in the price of the staple increases farm profits and hence full income.

$$\frac{dX_a}{dp_a}\delta p_a = \frac{\delta\pi}{\delta p_a}\delta p_a = Q\delta p_a$$

The profit effect equals output times the change in price.

When the price of a commodity e.g. maize increases, traditional demand theory suggests that the demand for maize will decline. However, for an agricultural household, the increase in the price of maize will result in an increase in household revenue (higher price times quantity of maize produced). This may increase the profit households realize from growing maize and hence increase the ability of some households to increase the quantity of maize they consume. Ultimately the direction of the response to the increased price of maize depends on the household characteristics and resource endowment. This is what this thesis has attempted to measure. An important point emphasized in Strauss, is that the presence of competitive product and factor markets is necessary for use of the assumption of separability.

The derivations above demonstrate that despite only focusing on supply response, we make an important assumption: All households maximize their utility in response to changing maize prices. This thesis has simply measured the observable characteristics of the households characteristics in relation to changing prices. Future research can go into more detail in attempting to measure the consumption related characteristics of rural households in order to do joint estimation of production versus consumption decision making.

A complete estimation of a non-separable household model requires consumption and production data for identical households (Strauss). This requirement is only partially met by the supplemental survey dataset. The information on household consumption patterns is not complete. At best we can measure apparent consumption using the standard definition ((production + purchases + gifts received)-(sales +stored quantity + gifts sent out). The price data must also have high variability. This is normally achieved through the inclusion of household level price data. However, the incorporation of naïve expectation theory in the relationship between price and production decisions resulted in the inclusion of lagged price data. As mentioned earlier, lagged household level price data is only available from the Post-harvest Survey datasets. In order to incorporate the PHS price data in this analysis, however, the estimates are aggregated to the district level resulting in a loss of variability at the household level.

The variable on household labour participation in the SS 2001 dataset is at the member level. This means that several members could have participated in off-farm labour activities. When the price of a commodity increases, the price rise is likely to have impacts on the household's ability to provide labour. Since our analysis is at household level, therefore we reclassify each household's participation in labour activities. We introduce a binary variable for household labour participation. Approximately 18.8 percent of all households interviewed in the SS01 survey reported having at least one member participating in off-farm labour activities. In addition, 40.4 percent of all household members who reported performing off-farm income did so on a smallholder farm or commercial farm, 19.7 percent were civil servants, and 16.4 percent did some non-agricultural type of work.

Findings from the HH04 survey reveal that 21.2 percent of all households reported at least one member performing off-farm labour activities. 38.5 percent of all members who reported off-farm labour activity worked on small holder and commercial farms, 17.9 percent were civil servants and 20.2 percent did some non-agricultural piece work. The above analysis highlights the fact that there are some structural rigidity that seem to limit the participation of households in off farm labour activities.

The data used for this analysis is based on rural households. There is considerably heterogeneity among rural farm households with significant numbers of households being net purchasers of staples. Another crucial assumption borrowed from Deaton is that goods and leisure are separable in preferences. The supplemental survey does not contain any variables to test whether or not this assumption is valid.

A household cannot consume more than it produces (own production, purchases from labour sales etc) unless it borrows or liquidates some assets. A thorough treatment (at the national level) of the interaction of shifts in demand and supply in response to price changes requires the estimation of three distinct models, a separable small and medium holding model, a separable consumption model and a separable large scale holding model. Price changes also influence the household's leisure labour mix. To what extent this is true is beyond the scope of this thesis (the data does not contain sufficient variables to render empirical analysis).

Given the various arguments made above, this thesis limits its analysis to a supply response model of rural households to changing maize prices. Consumption models are not included due to the limited nature of the data. In addition, the estimation of a complete demand system would require additional information on household consumption patterns. Anecdotal evidence from the FSRP maize value chain study (FSRP, 2009)[14] suggests that there is considerable flow of processed urban maize to rural centres during the lean period. This suggests a rural-urban interaction of demand that is not adequately addressed in the data[15].

Other Conceptual issues that are considered in the model design

Price Transmission

The Econometric model used in the analysis for this thesis relies on household farm gate prices. Therefore, no assumptions are necessary about the extent of price transmission from world to domestic markets as I measure the direct effect of local prices on smallholder behavior.

Zambia has experienced 10 deficit years out of the last 20. The Zambia annual food balance sheet details the projected annual demand for the main staple crops in the country. This demand is projected on the basis of annual demand for human

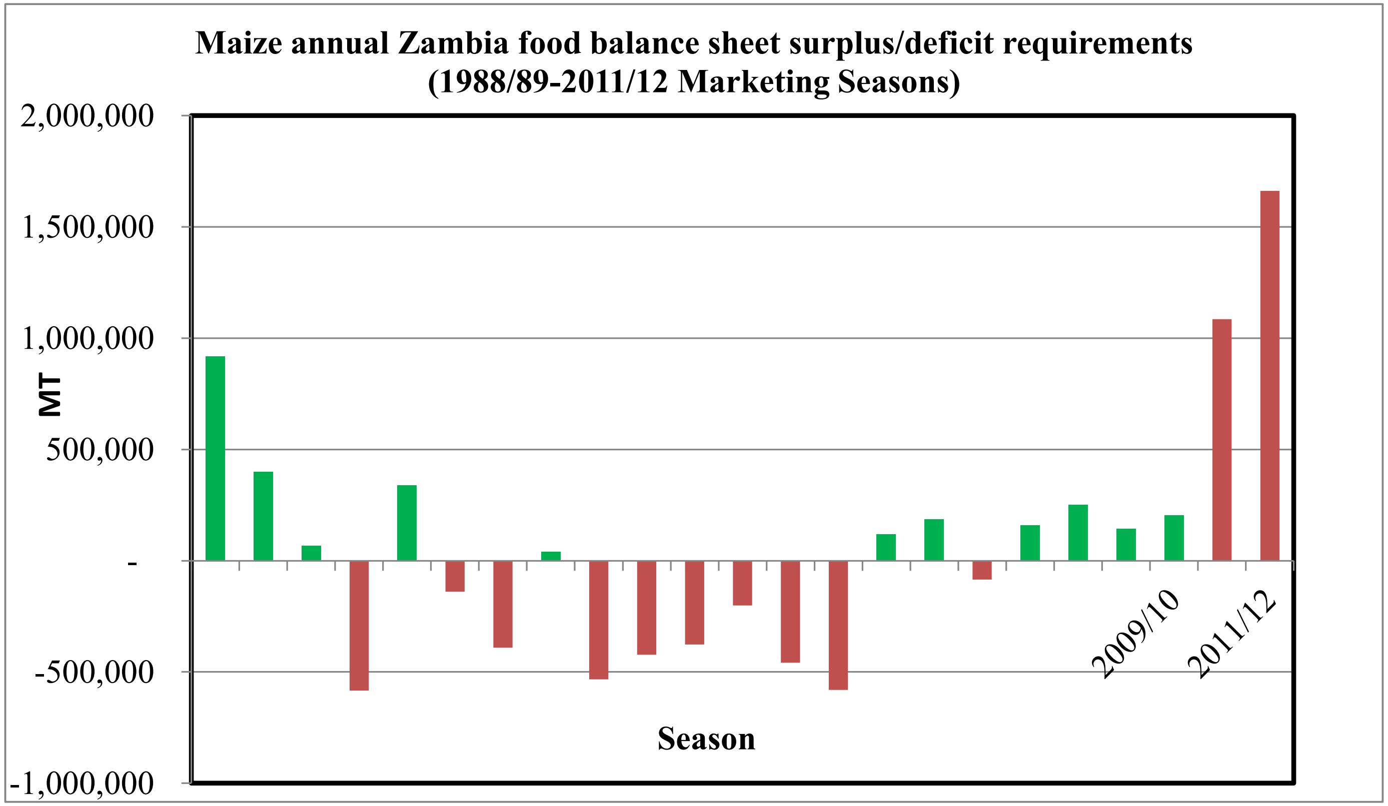
---

14 June/July 2009 survey of maize marketing in major growing areas. The survey also included several of the largest milling plants in the country. The major pattern that emerged from the survey was that from harvest up to about October/November, much of the maize consumed in the country is from the small and medium scale sector. However, from December to April, millers increasingly rely on stocks held by the large scale farming sector. Demand for processed mealie meal goes up, even into the rural areas.

15 The Post Harvest Surveys (up to 2004) and Supplemental survey have been based on a rural sample. The large scale farming sector which contributes significantly to national maize consumption during the lean period in Zambia is not included in the dataset

consumption. There is some level of substitution effects among the balance sheet crops depending on price and availability. If the price of maize increases relative to cassava, some consumers may increase consumption of cassava relative to maize. It is difficult to directly measure this substitution effect. However price signals are used to indirectly observe the elasticity of maize demand in relation to other crops. The substitution effect is captured in the model by the inclusion of the price of substitute crops to maize as independent variables.

Figure 6: Maize Surplus/Deficit requirements based on Zambia Annual Food Balance Sheets from '89 to '2011/12



Based on 1989/90 to 2011/12 Zambia Annual Food Balance Sheets

Government Policy

Government intervention in the Zambian agricultural sector has impacted producers and consumers;

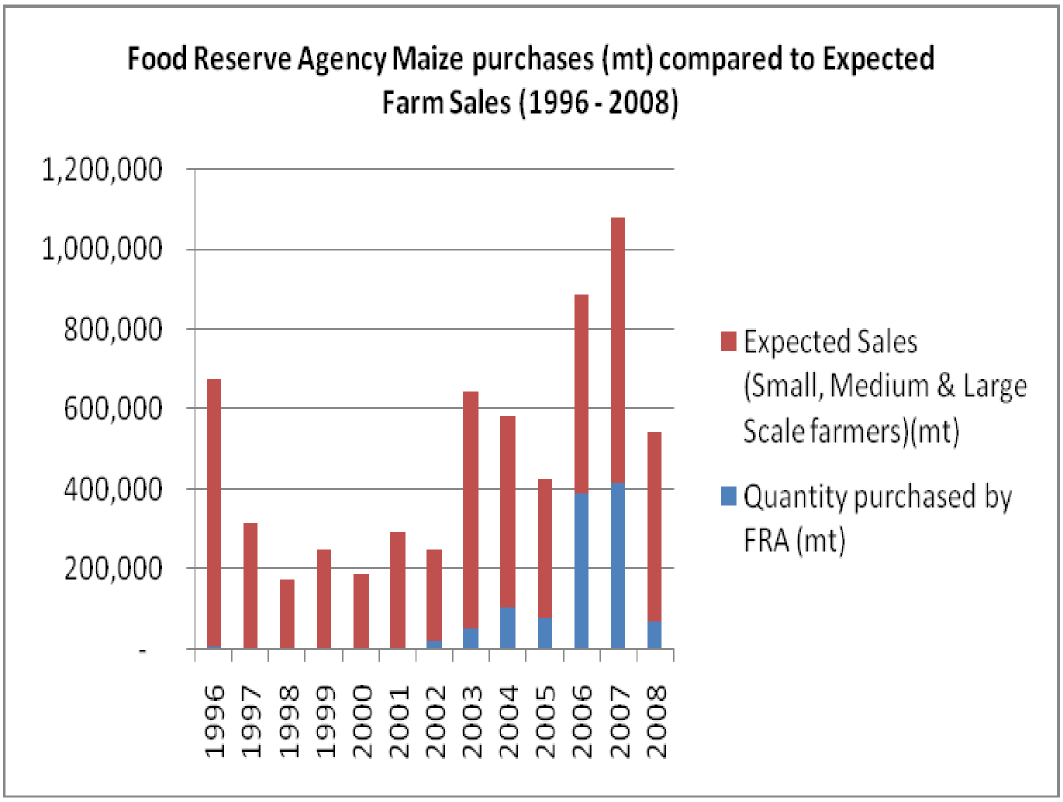
i) Producer impact of Government Policy

Since 2002/2003, the government has been implementing a policy to subsidize fertilizer and maize seed for small scale farmers. The presence of a subsidy may or may not impact a households decision to grow maize. The proportion of households using fertilizer has typically ranged between 31 – 36% (FSRP). Fertilizer was included as one of the variables in the inputs vector of the supply response model. Also included was a binary variable asking whether the household accessed Government subsidized inputs. This was interacted with the time variable. An important question is whether or not there is bias in the selection of recipients of the fertilizer.

A binary variable that asks whether the household sold maize to the FRA is included. This aims at capturing the impact of the state-run marketing channel on supply response. However, this variable may suffer from bias. Based on the focus group discussions of a maize value chain study, many farmers who tried to sell maize to the FRA complained that only well connected farmers were given priority to sell maize to the agency. Well-connected farmers are likely to be relatively well endowed with production assets and able to marshall resources from other means to invest in higher production[16]. To overcome this bias instrumental variables are used in the analysis.

[16] Several interviews with farmers who had taken their maize to the FRA sheds were held. The complaint of well connected farmers being given first priority was widespread.

Figure 7: FRA maize purchases since inception of the agency



In developing countries such as Zambia, regional price differences can also can provide useful insights on spatial supply response. A provincial dummy is included.

Price Expectation

Naïve price expectation assumes that the best forecast for a future price is a current price (Gomez, Love & Burton). This expectation does ignore the potential impact of changes in demand supply conditions on price. Although the predictive ability of naïve expectations can be poor, it is useful in situations where the collection of additional information is costly. This is typically the case for rural households. Rahji and Adewumi, in an analysis of the market supply response and demand for rice in Nigeria, used the expected price and area planted in the preceding year as a predictor of the area to be planted to rice in the current year in an analysis that assumed that farmers would not know with certainty what price they will receive at harvest. A naïve price expectation theory is used in this thesis. In Zambia, farmers plant their maize crop between November and the first week of January. It is assumed that the price of maize received by the farmer in the most recent season prior to the current planting season will have the most impact in influencing the household decision making. The 2006 price of maize received by the household is likely to have influenced the household decision making process for maize harvested in 2007. The large heterogeneity and spread of small holder farmers, coupled with distance to markets makes naïve expectation relatively reasonable for use in this model.

Econometric Methods Chosen

Fixed Effects modeling

In time series analysis, unobserved factors that influence our model over time are of two types, those that are constant and those that vary over time. Letting i denote the cross sectional unit and t the time period, Wooldridge writes a model with a single observed explanatory variable as;

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it} \quad t = 1,2,3. \qquad (1\text{-}9)$$

Relating the notation to the data in this thesis, i denotes the household and t the three time periods covered in the survey. Note that t does not change across i (households in our case). The intercept does change across time though. The error u_{it} is often called the idiosyncratic error or time-varying error, because it represents unobserved factors that change over time and affect y_{it}. These are very much like the errors in a straight time series regression equation (Wooldridge).

Fixed effects estimation models '(group dummies) control for group averages… because fixed effects models rely on intra-group action, you need repeated observations for each category, and a reasonable amount of variation of your key X variables within each category (Jayne). Consequently, the fluctuation in the rainfall data around a provincial mean is controlled for automatically. Using pooled OLS on the time series model without adjusting for the fixed effects would generate data that is inconsistent and biased.

An alternative to the fixed effect model is the Random effects model.

Wooldridge suggests that the Random effects is an attractive alternative to fixed effects under certain conditions; when we think the unobserved effect is uncorrelated with all the explanatory variables. If we have good controls in our equation, we might believe that any leftover neglected heterogeneity only induces serial correlation in the composite error term, but it does not cause correlation between the composite errors and the explanatory variables.

Fixed Effects versus Random effects

Equation (1-9) above becomes a random effects model when we assume that the unobserved effect ai is uncorrelated with each explanatory variable:

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it} \quad t = 1,2,3. \qquad (1\text{-}9)$$

$$\text{Cov}(xitj, ai) = 0, \; t = 1,2,\ldots, T; \; j = 1,2,\ldots, k.$$

Comparing the FE and RE estimates can be a test for whether there is correlation between the a_i and the x_{itj}, assuming that the idiosyncratic errors and explanatory variables are uncorrelated across all time periods. Both the Fixed and Random effects analysis results are presented.

One advantage of fixed effects analysis is that it does allow for attrition within the sample. This is an advantage if the attrition is correlated with the unobserved effect, a_{it}

Three household models are used. The first is an instrumental variables model with quantity of government fertilizer used by the household as the dependent variable. This initial model incorporates some instruments that are not correlated with area or yield but are correlated with access to government fertilizer. The resulting coefficients are then included in the two supply response models, one using area planted to maize and the other using maize yield rate as the dependant variables. Explanatory variables include household landholding size, price of maize in the previous season, cost of inputs in the current season, household asset base, household size, price of alternative crops, dummy variable representing advice received in previous season and Water Requirement Satisfaction Index, a variable that represents the percentage of plant rainfall requirements met for the yield model only.

Instrumental variables

As already stated, the inclusion of certain variables such whether the household sold maize to the FRA does introduce some bias into the model. This is because access to government services such as ability to sell maize to the FRA or access to government subsidies are probably connected to other criteria that are not included in the model. It is more likely that households with more social capital are likely to have easier access to Government subsidies than households without. This may result in correlation of some of the explanatory variables with the error term.

Instrumental Variable (IV) methods allow consistent estimation when the explanatory variables (covariates) are correlated with the error terms. This correlation may result from the dependent variable having a causal effect on at least one of the dependent variables and such variables have been from the model, or when the covariates are subject to measurement error. In this situation, ordinary linear regression generally produces biased and inconsistent estimates (Pearl, 2000). The instrumental variable would be correlated to the explanatory variables, conditional on the other covariates. As an example, access to Government fertilizer is correlated to whether the household has some social connections to the local agricultural authority who influence the decisions of who gets subsidized fertilizer and who does not. However, these social connections are not necessarily connected to the area that a household decides to plant to maize. Social connections can be an instrument in this particular example.

As an illustration, if we have the equation

$$Y_i = \beta_0 + \beta_1 X_i + u$$

Where
$$Y = yield$$
$$X = fertilizer\ used$$
$$u = error\ term$$

Any factors that influence a households access to fertilizer will be contained within the error term and

$$Cov\{x, u\} \neq 0$$

Under such conditions, using IV works to enable us obtain consistent estimators of β_0 *and* β_1. This is done through estimating a third variable that is correlated with X_i but not with Y_i. This principle is applied to the model on the household's acquisition of fertilizer.

$$Q_{fert} = \beta_0 + \beta_1 x_{rship} + \beta_2 x_{dist} + \beta_3 x_{fjob} + \beta_4 x_{fgroup} + \beta_5 x_{sexhead}$$

Where

Q_{fert} =quantity of fertilizer used by the household

$\beta_1 x_{rship}$ = relationship of head of household to headman (initial conditions)

$\beta_2 x_{dist}$ = distance to fertilizer markets

$\beta_3 x_{fjob}$= household member with a formal sector job

$\beta_4 x_{fgroup}$ = membership in a farmer group

These explanatory variables are likely to influence whether a household is able to access government fertilizer but are not directly correlated to Area planted to maize or maize yield.

The dependent variable coefficients obtained from the model are then used as an explanatory variable in the relevant household models. This operation is performed in STATA using the 'xtivreg' command.

Instrumental Variable Model Results

Table 9: OLS Model Instrumenting Quantity of Government Fertilizer acquired with Factors that are not directly correlated with Area planted to maize or Yield of Maize

	Quantity of Govt fertilizer Acquired		
	Supplemental Survey Year		
Explanatory Variables	2001	2004	2008
hh related to village head at time of land allocation	-9.378*	-5.676**	-5.537**
	(5.693)	(2.347)	(2.583)
number of hh members in formal employment	21.50	21.30**	9.701
	(21.21)	(10.31)	(6.568)
Does hh purchase inputs with a group	164.6***	17.98	8.826
	(48.73)	(11.08)	(8.120)
distance of fertilizer access point	1.153***	0.443***	1.633***
	(0.301)	(0.169)	(0.340)
Constant	25.42***	25.31***	51.69***
	(4.933)	(3.700)	(5.926)
Observations	4281	4281	4281
R-squared	0.018	0.006	0.013
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Household Models

Two small and medium scale household supply response models are generated with quantity of maize produced, area under maize and yield rate for maize as the dependant variables in each of these models respectively. Explanatory variables include, household land holding size, price of maize in previous season, cost of inputs in current season, household asset base, household size, price of alternative crops, dummy variable representing advice received in previous season and a variable representing rainfall.

Model 1-Area planted to maize as dependent variable

$$S_m = \beta_0 + \beta_1 x_{hh\_land} + \beta_2 x_{pr\_mz\_1} + \beta_3 x_{cst\_inp} + \beta_4 x_{hhasset} + \beta_5 x_{hhsize} + \beta_6 x_{pr\_alt} + \beta_7 x_{dd\_adv} + \beta_8 x_{rain} + \beta_9 \hat{Q}_{govt\_fert}$$

where

$S_m =$ **Area planted to maize by small and medium scale household,**

$\beta_1 x_{hh\_land}$ **= size of household land,**

$\beta_2 x_{pr\_mz\_1}$ **= expected price of maize (based on deflated price of maize in previous season),**

$\beta_3 x_{cst\_inp}$ **= cost of inputs (cost of fertilizer is used as proxy),**

$\beta_4 x_{hhasset}$ **= Vector of household assets,**

$\beta_5 x_{hhsize}$ **= household size (adult equivalent),**

$\beta_6 x_{pr\_alt}$ **= price of alternative crops (groundnuts),**

$\beta_7 x_{dd\_adv}$ **= dummy variable for extension advice**

$\beta_9 \hat{Q}_{govt\_fert}$ **= coefficients of quantity of Government fertilizer used based on IV model**

Model 2-Yield of maize as dependent variable

$$S_y = \beta_0 + \beta_1 x_{\ln mz\_df} + \beta_2 \hat{Q}_{govt\_fert} + \beta_3 X_{assets} +$$
$$\beta_4 x_{hhsize} + \beta_6 x_{time} + \beta_7 x_{soiltype} + \beta_8 x_{rain}$$

where

S_y **= Yield rate of maize by small and medium scale household,**

$\beta_1 x_{\ln mz\_df}$ **= expected price of maize (based on deflated price of maize in previous season),**

$\beta_2 \hat{Q}_{govt\_fert}$ **= quantity of Government fertilizer used,**

$\beta_3 X_{assets}$ **= Vector of household assets,**

$\beta_5 x_{hhsize}$ **= household size (adult equivalent),**

$\beta_6 x_{time}$ **= technological progress represented by time dummy,**

$\beta_7 x_{soiltype}$ **=soil type,**

$\beta_8 x_{rain}$ **= adequacy of rainfall received,**

$\beta_9 \hat{Q}_{govt\_fert}$ **= coefficients of quantity of Government fertilizer used based on IV model**

The two supply response models estimate the total national response both in terms of area and yield rates to changes in the price of maize. The results from the yield rate model are particularly interesting because they have potential policy implications on adaptive mechanisms adopted by land constrained households

Econometric Results Section

Regression results are presented in tables 10 through table 17. Regression coefficients and marginal effects of the area and yield models are presented. Also included with the parameter estimates are the standard errors and statistics on levels of significance. Other statistics are also included.

The Econometric analysis was done using Stata version 9. The initial variable manipulation was done using SPSS version 17 and the resulting datasets transferred to STATA. The main research questions analyzed in this thesis include the supply response of maize farmers to changes in maize prices in Zambia among small holder rural households. Emphasis is placed on measuring the differential supply response rates among smallholder farmers given the great heterogeneity within the small and medium scale sector. The second broad objective is to compute elasticities to measure the relationship between maize prices as well as other variables of the production function on area and yield rates for maize.

The analysis in this section of the report restricts itself to households who grew maize at least once in any of the three survey years. Households that did not grow maize at all were dropped from the analysis. All models included provincial binary variables for all provinces in Zambia except Lusaka Province which lacks a major smallholder presence relative to other provinces due to its limited geographical size and disproportionate urbanization levels.

Regression Diagnostics

Two regression diagnostics models were run to check for attrition bias and to check for multi-collinearity. The 'collin' function in STATA was used to check for multi-collinearity. The maize price produced a VIF slightly above the recommended thresh-hold of 10. This is due to the probable correlation between maize prices and groundnut prices, the other alternative crop price included in the model. The variables for age of household head and age of head interacted with price had a relatively high VIF but are maintained in the model.

Area Models

Two alternative supply response models using area as the dependent variable are presented, both using fixed effects. The initial base model is generated without price interactions and acts to check for robustness. A log-linear specification was used for most of the variables. However, continuous variables such as value of assets are expressed in log terms. The area models measure the relationship between price and other variables and decisions relating to changes in quantity as measured by area under maize production.

The overall model is statistically significant with a probability > F = 0.0000 and an overall

R-squared = 0.2445 for the base model without price interaction and an overall R-squared = 0.2729 for the model with price interaction.

The overall direction of the relationship between area and most of the explanatory variables is as we would expect. In the area model with price interaction, area planted maize is positively related to the quantity of subsidized Government fertilizer accessed by households and negatively related to the full market price of fertilizer. However, the coefficient of Government fertilizer is not significant even at the 10% level and the magnitude of the economic relationship is very small. If this relationship does indeed hold, this could be due to the limited amount of fertilizer that households are able to access relative to total fertilizer used in the country. Also whether or not a household will actually access Government fertilizer is often only known very late during the planting season and the households may have already made planting decisions by the time they are actually able to access the Government fertilizer. In all the survey years approximately 33% of all households using Government fertilizer did not receive it on time. In 1999 alone when the planting decisions for the 1999/2000 season were being made, up to 54% of households never received the fertilizer on time. A 1% increase in the price of fertilizer corresponds to a -0.00119% reduction in the area planted to maize. This result is significant at the 1% level.

As expected the price of maize is positively related to area planted to maize. Both variables are in logs so this result represents an elasticity. A 1% increase in the expected price of maize increases area planted to maize by 0.66%. This result is significant at the 5% level. The price of groundnuts is negatively correlated to area planted to maize as expected. However, this result is not statistically significant.

Binary variables representing the quintiles of total land area held by households were created for each of the time periods between 2001 and 2004 as well as between 2004 and 2008. All the results are both statistically significant at the 1% significance level and economically significant. The first and second quintiles of landholding had a negative relationship with area planted to maize all things being equal. However, the fourth and fifth quintile had a positive relationship. For a given level of all other variables, households with larger land area are able to respond to increases in maize prices by increasing area under maize. However, the magnitude of the relationship was different between the different survey periods. Households in the first quintile maintained a consistent reduction in area planted to maize between the two 'jumps'. However, households in the second quintile reduced their area under maize by a small margin in 2008 compared to 2004. Households in the fifth quintile increased area planted to maize by 0.382 and 0.448 between the 2001-2004 period and 2004-2008 period respectively. These results are consistent with conventional wisdom among most practitioners in the agriculture sector in Zambia. Households with more access to land are more likely to respond to changes in maize production compared to households with more limited access to land. In addition, the household land area variable interacted with maize prices is statistically significant at the 1% level and positively related to maize prices.

The reasons for the negative relationship between area planted to maize and maize prices could be due to the fact that significant numbers of small holder households are net maize buyers. They do not produce sufficient quantities to feed their families. Consumption requirements are often met via cash purchases, livestock sales and even payment in kind as farm labour on larger holdings. Higher prices of maize would therefore translate into longer hours devoted to labour in order to earn enough money for an equivalent quantity of maize, leaving less time to work on own fields.

Three variables representing education levels of household members were included in the model. These are highest level of education level of head of household, spouse and other members of the household. All three variables were positively economically related to area planted to maize. However, only the 'education of other household member' was statistically significant at the 10% level. Interestingly, the coefficient of education level of spouse was larger than that of the head of household and the coefficient of education level of other members was the highest of the three. Similar studies, although focusing on livestock marketing have found that households with higher education levels tend to participate more in the market (Ehui, 2003) and (Holloway, 2000).

The provincial binary variables produce very interesting results. Initially, the supply response to area was negative for households in Central province for the 2001-2004 interval. However, the 2004-2008 interval produced a positive supply response. However, both coefficients were not statistically significant. The Copperbelt province had negative coefficients for both periods and these were statistically significant at the 5% level. The coefficients of most of the provincial binary variables roughly correspond to availability of land. Eastern, Southern and Copperbelt have relatively less available land to facilitate expansion of crops. The coefficients for these provinces are negative. Luapula, Northern, North-western and Western provinces have relatively more available land and all four have positive coefficients for the 2008 period with levels of significance at the 10%, 5%, 10% and 1% level respectively.

Yield Model

The yield model includes variables that relate to technological and management changes to production as well as exogenous variables such as weather and soil type. The yield model is also expressed in log linear format with some continuous variables such as value of assets expressed in log form.

The overall significance of the yield model is relatively low compared to the area model. This is explained by the difficulty in obtaining rainfall that adequately models rainfall performance in Zambia. Currently rainfall data series is based on approximately 40 reporting stations out of the 72 districts in the country.

The most significant result from the yield model is that the proxy for the availability of extension advice on use of crop residues was positively related to yield and significant at the 5% level. In addition, the highest education level of the spouse and other household members was positively related to yield. However, only the education level for 'other members' was significant at the 5% level. This is expected because in most rural households, much of the actual work in land husbandry is actually done by the spouse and other household members. Better knowledge of agricultural practices by the spouse and members can tend to produce better yields.

The provincial binary variables for Luapula, Northern and North Western provinces for the 2004-2008 periods were positively related to yield and statistically significant at the 1%, 1% and 10% levels respectively.

Table 10: Area Model

VARIABLES	Coefficient lnmz_area		Standard Errors
quantity of fert accessed through government channels	0.000170		0.000659
deflated log price of maize	0.663	**	0.299
deflated log price of groundnut	-0.216		0.248
real price of fertilizer per kg	-0.00119	***	0.000412
sex of hh head	-0.0827	*	0.0443
Quartile 1 area dummy 2004	-0.568	***	0.0453
Quartile 1 area dummy 2008	-0.567	***	0.0470
Quartile 2 area dummy 2004	-0.236	***	0.0379
Quartile 2 area dummy 2008	-0.188	***	0.0403
Quartile 4 area dummy 2004	0.195	***	0.0370
Quartile 4 area dummy 2008	0.191	***	0.0381
Quartile 5 area dummy 2004	0.382	***	0.0457
Quartile 5 area dummy 2008	0.448	***	0.0573
Adult Equivalent per hh	0.0224	***	0.00575
log of value of hh productive assets	1.66e-09		2.43e-09
total cattle raised by hh	0.00125		0.000836
total pigs raised by hh	0.00407	**	0.00169
area interacted with maize price	0.000307	***	4.18e-05
age of household head	0.0175	***	0.00307
age of head interacted with deflated maize price	-0.0243	***	0.00449
dependency ratio- count of hh members less than 15 and over 60 by hh size	-0.0323		0.0372
education level of head	0.00565		0.00563
education level of spouse	0.00853		0.00695

Table 10 (cont'd)

education level of highest educated other member	0.00935	*	0.00491
education level of head interacted with deflated maize price	-0.00234		0.00310
education level of spouse interacted with deflated maize price	-0.00137		0.00352
education level of other members interacted with deflated maize price	-0.00596	**	0.00269
dummy central prov 2004	-0.0250		0.0595
dummy central prov 2008	0.0730		0.0713
dummy copperbelt prov 2004	-0.233	**	0.107
dummy copperbelt prov 2008	-0.179	**	0.0865
dummy eastern prov 2004	-0.303	***	0.0950
dummy eastern prov 2008	-0.340	**	0.141
dummy luapula prov 2004	0.0616		0.141
dummy luapula prov 2008	0.203	*	0.104
dummy northern prov 2004	-0.253	**	0.106
dummy northern prov 2008	0.176	**	0.0741
dummy north western prov 2004	0.236	*	0.125
dummy north western prov 2008	0.271	*	0.139
dummy southern prov 2004	-0.283	**	0.128
dummy southern prov 2008	-0.100		0.175
dummy western prov 2004	0.155		0.157
dummy western prov 2008	0.400	***	0.0753
Constant	-1.946		1.420
Observations	10118		
Number of hhid	3912		

Source: Estimated from the 2001, 2004 & 2008 Supplemental Survey to the 1999/2000 Post-harvest Survey of small and medium sized holdings Notes: Details of the explanatory variables are given in the Data Section. The dependent variable is a continuous variable of area planted to maize converted to logs. Only households that grew maize in at least one survey year were used in the analysis. All coefficients indicate statistical significance at the 10*, 5** and 1*** percent levels, respectively

Table 11: Marginal Effects Area Elasticities

VARIABLES	lnmz_area Coefficients		Standard Errors
quantity of fert accessed through government channels	0.00792		0.0307
real price of fertilizer per kg	-0.222	***	0.0765
sex of hh head	-0.0986	*	0.0528
Quartile 1 area dummy 2004	-0.0262	***	0.00209
Quartile 1 area dummy 2008	-0.0232	***	0.00192
Quartile 2 area dummy 2004	-0.0167	***	0.00268
Quartile 2 area dummy 2008	-0.0123	***	0.00264
Quartile 4 area dummy 2004	0.0145	***	0.00275
Quartile 4 area dummy 2008	0.0154	***	0.00307
Quartile 5 area dummy 2004	0.0202	***	0.00241
Quartile 5 area dummy 2008	0.0397	***	0.00507
Adult Equivalent per hh	0.119	***	0.0305
total cattle raised by hh	0.00391		0.00262
total pigs raised by hh	0.00805	**	0.00335
area interacted with maize price	0.135	***	0.0184
age of household head	0.659	***	0.116
age of head interacted with deflated maize price	-0.562	***	0.104
dependency ratio- count of hh members less than 15 and over 60 by hh size	-0.00939		0.0108
education level of head	0.0299		0.0299
education level of spouse	0.0279		0.0227
education level of highest educated other member	0.0477	*	0.0251
education level of head interacted with deflated maize price	-0.0163		0.0216

Table 11 (cont'd)

education level of spouse interacted with deflated maize price	-0.00581		0.0149
education level of other members interacted with deflated maize price	-0.0427	**	0.0193
dummy central prov 2004	-0.00100		0.00239
dummy central prov 2008	0.00336		0.00328
dummy copperbelt prov 2004	-0.00510	**	0.00234
dummy copperbelt prov 2008	-0.00419	**	0.00202
dummy eastern prov 2004	-0.0263	***	0.00826
dummy eastern prov 2008	-0.0321	**	0.0133
dummy luapula prov 2004	0.000731		0.00168
dummy luapula prov 2008	0.00305	*	0.00156
dummy northern prov 2004	-0.0118	**	0.00495
dummy northern prov 2008	0.00879	**	0.00369
dummy north western prov 2004	0.00475	*	0.00252
dummy north western prov 2008	0.00563	*	0.00287
dummy southern prov 2004	-0.0137	**	0.00619
dummy southern prov 2008	-0.00529		0.00923
dummy western prov 2004	0.00552		0.00560
dummy western prov 2008	0.0151	***	0.00284
Observations	10118		
Number of groups	3912		

Source: Marginal effects of coefficients on Area without price interaction. Coefficients represent the percentage change in Area planted resulting from a 1% change in coefficient. All coefficients indicate statistical significance at the 10*, 5** and 1*** percent levels, respectively. Coefficients of explanatory variables that are already in logs are not included as they already represent elasticities with respect to area.

Table 12: Yield Model

VARIABLES	Coefficient lnmz_yield		Standard Errors
quantity of fert accessed through government channels	-0.000927		0.00259
deflated log price of maize	0.898		1.201
deflated log price of groundnut	-2.235		2.617
proportion of hh in SEA reporting timely top availability	0.000818		0.000803
sex of hh head	-0.201		0.149
real price of fertilizer per kg	0.00227		0.00143
Quartile 1 area dummy 2004	-0.832	***	0.154
Quartile 1 area dummy 2008	-0.778	***	0.155
Quartile 2 area dummy 2004	-0.339	**	0.135
Quartile 2 area dummy 2008	-0.108		0.140
Quartile 4 area dummy 2004	-0.0929		0.134
Quartile 4 area dummy 2008	0.0272		0.137
Quartile 5 area dummy 2004	-0.281	*	0.167
Quartile 5 area dummy 2008	-0.0631		0.216
Adult Equivalent per hh	0.0246		0.0205
log of value of hh productive assets	4.66e-10		9.72e-09
area interacted with maize price	0.000259	*	0.000153
was use of crop residues advice available to hh?	0.00305	**	0.00124
total cattle raised by hh	0.00314		0.00314
age of household head	0.00639		0.00456
education level of head	-0.00786		0.0192
education level of spouse	0.00769		0.0236
education level of highest educated other member	0.0325	**	0.0163
education level of head interacted with deflated maize price	-0.00302		0.0104
educ level of spouse interacted with deflated maize price	-0.0112		0.0119
education level of other members interacted with deflated maize price	-0.00760		0.00890
dependency ratio- count of hh members less than 15 and over 60 by hh size	-0.186		0.128
minimum rainfall 1990-91	0.000411		0.00188
maximum rainfall 1990-91	-0.000276		0.00185
dummy central prov 2004	-0.0206		0.696
dummy central prov 2008	0.809	*	0.445
dummy copperbelt prov 2004	0.0115		0.422

Table 12 (cont'd)

dummy copperbelt prov 2008	0.428		0.333
dummy eastern prov 2004	-0.519		0.874
dummy eastern prov 2008	-0.302		0.806
dummy luapula prov 2004	0.478		0.474
dummy luapula prov 2008	1.618	***	0.423
dummy northern prov 2004	0.494		0.399
dummy northern prov 2008	1.720	***	0.312
dummy north western prov 2004	2.435		1.790
dummy north western prov 2008	2.459	*	1.489
dummy southern prov 2004	-0.244		1.183
dummy southern prov 2008	1.275		1.899
dummy western prov 2004	-0.781		0.843
dummy western prov 2008	0.566		0.480
acriso04	0.349		0.245
acriso08	0.0162		0.238
alisol04	0.384		0.402
alisol08	-0.185		0.390
arenos04	-0.235		0.274
arenos08	-0.156		0.265
cambis04	-0.547		0.796
cambis08	0.266		0.748
ferral04	0.208		0.370
ferral08	0.122		0.346
fluvis04	0.163		0.974
fluvis08	0.342		0.894
gleyso04	1.104	**	0.458
gleyso08	0.714	*	0.397
histos04	3.028	***	0.474
histos08	0.429		0.454
leptos04	0.372		0.277
leptos08	-0.0264		0.365
lixiso04	0.203		0.325
lixiso08	-0.274		0.310
luviso04	0.587		0.410
luviso08	0.243		0.396

Table 12 (cont'd)

phaeoz04	0.873		0.868
phaeoz08	0.668		0.834
planos04	-2.019	*	1.031
planos08	-0.720		0.915
podzol04	0.458		0.451
podzol08	0.349		0.436
regoso04	0.317		1.018
regoso08	-1.350		0.889
solone04	0.108		0.464
solone08	-0.0643		0.453
vertis04	0.0626		0.356
vertis08	-0.630	*	0.340
Constant	12.91		14.75
Observations	11264		
Number of hhid	3932		

Source: Estimated from the 2001, 2004 & 2008 Supplemental Survey to the 1999/2000 Post-harvest Survey of small and medium sized holdings Notes: Details of the explanatory variables are given in the Data Section. The dependent variable is a continuous variable of maize yield converted to logs. Only households that grew maize in at least one survey year were used in the analysis. All coefficients indicate statistical significance at the 10*, 5** and 1*** percent levels, respectively

Table 13: Marginal Effects Yield Elasticities

VARIABLES	lnmz_yield Coefficients		Standa rd Errors
quantity of fert accessed through government channels	-0.0383		0.107
proportion of hh in SEA reporting timely top availability	0.0499		0.0490
sex of hh head	-0.241		0.179
real price of fertilizer per kg	0.423		0.268
Quartile 1 area dummy 2004	-0.0438	***	0.00813
Quartile 1 area dummy 2008	-0.0371	***	0.00740
Quartile 2 area dummy 2004	-0.0233	**	0.00927
Quartile 2 area dummy 2008	-0.00709		0.00921
Quartile 4 area dummy 2004	-0.00634		0.00918
Quartile 4 area dummy 2008	0.00207		0.0104
Quartile 5 area dummy 2004	-0.0132	*	0.00783
Quartile 5 area dummy 2008	-0.00518		0.0177
Adult Equivalent per hh	0.128		0.107
log of value of hh productive assets	0.000565		0.0118
area interacted with maize price	0.110	*	0.0653
was use of crop residues advice available to hh?	0.135	**	0.0548
total cattle raised by hh	0.00905		0.00906
age of household head	0.239		0.171
education level of head	-0.0415		0.101
education level of spouse	0.0248		0.0763
education level of highest educated other member	0.163	**	0.0818
education level of head interacted with deflated maize price	-0.0209		0.0723
education level of spouse interacted with deflated maize price	-0.0468		0.0497
education level of other members interacted with deflated maize price	-0.0533		0.0623
dependency ratio- count of hh members less than 15 and over 60 by hh size	-0.0537		0.0370
minimum rainfall 1990-91	0.376		1.719
maximum rainfall 1990-91	-0.303		2.033
dummy central prov 2004	-0.000781		0.0264
dummy central prov 2008	0.0344	*	0.0189
dummy copperbelt prov 2004	0.000222		0.00813
dummy copperbelt prov 2008	0.00973		0.00757
dummy eastern prov 2004	-0.0396		0.0666
dummy eastern prov 2008	-0.0258		0.0687
dummy luapula prov 2004	0.00810		0.00803

Table 13 (cont'd)

dummy luapula prov 2008	0.0351	***	0.00917
dummy northern prov 2004	0.0250		0.0202
dummy northern prov 2008	0.0948	***	0.0172
dummy north western prov 2004	0.0452		0.0332
dummy north western prov 2008	0.0544	*	0.0329
dummy southern prov 2004	-0.0110		0.0532
dummy southern prov 2008	0.0615		0.0916
dummy western prov 2004	-0.0271		0.0293
dummy western prov 2008	0.0206		0.0175
acriso04	0.0430		0.0302
acriso08	0.00228		0.0335
alisol04	0.00296		0.00310
alisol08	-0.00154		0.00325
arenos04	-0.00473		0.00552
arenos08	-0.00313		0.00532
cambis04	-0.000728		0.00106
cambis08	0.000402		0.00113
ferral04	0.00236		0.00420
ferral08	0.00150		0.00424
fluvis04	0.000144		0.00086 5
fluvis08	0.000395		0.00103
gleyso04	0.00804	**	0.00333
gleyso08	0.00526	*	0.00293
histos04	0.0151	***	0.00235
histos08	0.00267		0.00282
leptos04	0.0153		0.0114
leptos08	-0.00119		0.0165
lixiso04	0.00473		0.00757
lixiso08	-0.00698		0.00790
luviso04	0.00428		0.00298
luviso08	0.00184		0.00299
phaeoz04	0.00101		0.00100
phaeoz08	0.000890		0.00111

Table 13 (cont'd)

planos04	-0.00179	*	0.000915
planos08	-0.000831		0.00106
podzol04	0.00232		0.00228
podzol08	0.00192		0.00240
regoso04	0.000254		0.000813
regoso08	-0.00156		0.00103
solone04	0.000544		0.00235
solone08	-0.000360		0.00253
vertis04	0.000679		0.00385
vertis08	-0.00749	*	0.00405
Observations	11264		
Number of groups	3932		

Source: Marginal effects of coefficients on Yield model with price interaction. Coefficients represent the percentage change in yield resulting from a 1% change in coefficient. All coefficients indicate statistical significance at the 10*, 5** and 1*** percent levels, respectively. Coefficients of explanatory variables that are already in logs are not included as they already represent elasticities with respect to yield.

CONCLUSIONS AND IMPLICATIONS

Higher global average prices for maize present both a threat and opportunity to agricultural households in Zambia. This study attempted to analyze the economic relationship between maize prices, prices of alternative crops as well as other factors of production on the supply response of small holder farmers. Nationally representative empirical data measuring changes in household agricultural variables over a period of 8 years was used in the study.

The thesis studied two questions, the nature of the relationship between factors of production and area planted to maize as well as the distributional effects of maize price changes on households with different land holding sizes. The conclusion and implications are discussed accordingly.

Changes in area planted to maize were found to be influenced by several important factors; increases in the price of maize had a negative impact on households with smaller land holding sizes. In general households that are not self-sufficient in maize production will be made worse off by higher maize prices. In the case of households that rely on labour sales to meet purchase requirements for maize, higher prices will directly translate into longer hours of work in order to meet commensurate quantities of maize.

Households with higher land holdings tend to be households with larger asset holdings as well. They are able to increase deployment of these assets to meet the opportunity presented by higher prices, especially if they are self-sufficient in maize production. Subsidized fertilizer does not seem to have the impact increasing area under maize as anticipated. This would suggest that supply side policies such as input subsidies would not have the desired impact on national output if not properly managed. This would involve more timely availability of inputs etc. Labour availability is crucial to supply response. The adult equivalent variable is positive, economically significant and statistically significant at the 1% level. Attempts to stimulate maize production would therefore be limited by labour constraints, in addition to the land constraints faced by some households.

A significant proportion of all maize growing households cultivate less than 5 hectares of maize. The households with the most scope for responding to higher prices lie at the top of the quintile distribution. However, such households are relatively fewer than the typical small-scale growing households.

There are considerable differences in supply response across provinces. In Zambia, Luapula, Northern, North Western and Western provinces have relatively larger land holding sizes compared to other parts of the country. There is relatively more available land in these provinces. However, these provinces also lie further away from the major markets in the country. They also have farming systems that grow relatively more cassava compared to other parts of the country. Stimulating the maize value chain in these provinces requires several innovative policy proposals.

It is not clear whether the very low economic relationship between quantity of Government fertilizer available and area planted to maize is due to model mis-specification or whether it actually does hold. However, the relatively low impact of Government fertilizer on maize planting decisions should result in some policy review. Currently, the Zambian Government is spending over $100 million dollars a year, to subsidize fertilizer usage for 20% of the area planted to maize in the country. Studies like this call into question the benefits of such a colossal investment given all the other major priorities in the sector such as investment in research and extension advice.

The statistically significant relationship between availability of extension advice on usage of crop residue (a proxy for conservation farming) points to the fact investment in better research and extension delivery methods will yield even more dividends than merely providing blanket subsidies to farmers. Even more telling is the fact that the coefficient of availability of extension advice on yield is both statistically significant at the 5% level and economically more important that the availability of fertilizer.

Implications for Future Research

A major draw-back in both the yield models is the fact that rainfall data in Zambia is not sufficiently captured spatially in order to improve the analytical process. As already stated only about 40 district level data collection points exist in the country out of 72 districts. A single rainfall observation point per district is not even sufficient to measure the economic significance of rainfall variations on yield performance. More investment into rainfall data collection needs to be made.

In order to have a well measured economic analysis of the relationship between prices and supply response, it is necessary to improve the collection process for price data in Zambia. Collection of farm-gate prices needs to be institutionalized into the two main annual surveys conducted by CSO and MACO, namely the Crop Forecast and Post-Harvest Surveys. Monthly sales prices should be collected in order for future price analysis work to be improved upon.

It is necessary for a complete model to measure demand and supply systems to be estimated. Future work in this field should try to use the supplemental survey dataset to measure the net effect of consumption and production decisions given changing maize prices.

APPENDICES

APPENDIX A: DESCRIPTIVE STATISTICS OF MAIN EXPLANATORY VARIABLES USED IN REGRESSION ANALYSIS

Table 14: Area planted to maize (2001, 2004 & 2008 Supplemental Surveys)

		Area planted to maize (ha)			
Year	(N)	Minimum	Maximum	Mean	Std. Deviation
1999	7539	.00	17.00	.93	1.47
2002	5381	.00	26.33	.80	1.24
2006	6378	.02	51.00	1.34	2.04

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 15: Household Maize harvest (mt) (2001, 2004 & 2008 Supplemental Surveys)

	Household Maize harvest (mt)				
Year	(N)	Minimum	Maximum	Mean	Std. Deviation
1999	7417	.00	60.30	1.3820	3.00529
2002	5381	.00	89.13	1.3814	3.46059
2006	6371	.00	105.80	2.3953	5.56616

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 16: Household Maize yield (2001, 2004 & 2008 Supplemental Surveys)

	Household Maize yield (mt/ha)				
Year	(N)	Minimum	Maximum	Mean	Std. Deviation
1999	6142	.00	9.20	1.4896	1.26448
2002	4621	.00	9.20	1.5962	1.23841
2006	6654	.00	9.94	1.6124	1.26881

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 17: Area under maize as a proportion of total area under crops and fallow (2001, 2004 & 2008 Supplemental Survey)

	N	Mean Proportion (%)	Std. Deviation
1999	7579	42.40	36.74
2002	5315	44.67	33.15
2006	7787	43.68	33.18

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 18: Mean maize yields (2001, 2004 & 2008 Supplemental Survey)

Year	Mean maize yields				
	Percentile Group of area				
	1	2	3	4	5
1999	1410.84	1491.20	1540.07	1493.72	1501.98
2002	1307.15	1485.95	1589.02	1713.07	2005.83
2006	1457.73	1478.10	1544.29	1570.41	1887.82

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 19: Mean Yield of Maize by Sex of Household Head, Year & Quintile of Land Held by Household

Year			Mean maize yield (kgs/ha)					
			Quintile Ranking of Total Land Area Held by Household (Crop & Fallow)					
			1	2	3	4	5	Total
1999	sex of hh head	Male	1671.51	1707.60	1702.61	1630.53	1587.08	1649.53
		Female	1331.17	1341.86	1359.64	1346.81	1391.08	1350.19
2002	sex of hh head	Male	1394.24	1538.57	1653.00	1739.81	2002.83	1665.64
		Female	1157.48	1359.78	1324.86	1636.23	1941.15	1350.40
2006	sex of hh head	Male	1511.97	1559.89	1554.97	1607.90	1909.57	1666.27
		Female	1356.21	1278.33	1509.41	1385.95	1685.62	1413.84

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 20: Total Quantity of Fertilizer used by Quintile of Area held by Household

Year	Mean quantity of fertilizer applied by hh (kgs)				
	Quintile Ranking of Area Held by Household				
	1	2	3	4	5
1999	179.24	213.48	234.06	290.89	639.43
2002	124.65	139.41	193.77	279.26	677.21
2006	125.63	184.17	220.37	308.52	841.46

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 21: Total fertilizer used by household, by sex of head of household, by quintile ranking of area held by household

Year			Mean quantity of fertilizer applied by hh (kgs)				
			Quintile Ranking of Area Held by Household				
			1	2	3	4	5
1999	sex of hh head	Male	177.74	227.62	241.75	295.63	654.43
		Female	190.28	138.46	189.59	255.14	499.11
2002	sex of hh head	Male	125.06	142.40	192.42	285.85	677.06
		Female	141.93	126.85	172.05	203.25	641.76
2006	sex of hh head	Male	129.97	190.36	220.10	314.51	857.27
		Female	107.68	158.71	221.40	270.76	696.04

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 22:Partial Household-Level and National Maize Fertiliser Information: National Household-Level Net Yearly Income and Related Food Security Categorization, 2007/2008, 2003/2004

Seller type	Households			Households Producing Maize		Households Producing Maize With Fertiliser	
	Tercile	% of HHs	Number of HHs	Number of HHs	% of HHs	Number of HHs	% of HHs
	Low	33.3	497,736	371,878	74.7	82,070	16.5
2007/2008 S. Survey	Med	33.3	497,697	414,671	83.3	132,152	26.6
	High	33.3	497,764	445,688	89.5	252,546	50.7
National HH – Level Net Income	**Ave.**	**100**	**1,493,197**	**1,232,237**	**82.5**	**466,768**	**31.3**
	Low	33.3	414,608	306,883	74	63,760	15.4
2003/2004 S. Survey	Med	33.4	415,003	332,419	80.1	90,158	21.7
	High	33.3	414,200	373,526	90.2	200,539	48.4
National HH – Level Net Income	**Ave.**	**100**	**1,243,811**	**1,012,827**	**81.4**	**354,457**	**28.5**
	Low	33.3	369,933	235,180	63.6	34,357	9.3
2000/2001 S. Survey	Med	33.3	370,128	283,228	76.5	55,819	15.1
	High	33.3	369,836	322,480	87.2	136,035	36.8
National HH – Level Net Income	**Ave.**	**100**	**1,109,898**	**827,372**	**76.5**	**226,211**	**20.4**

Source: Supplemental Surveys to the 1999/2000 Post Harvest Survey, Central Statistical Office, 2007/2008, 2003/2004 and 2000/2001 Marketing Seasons. Productive assets in 2007/2008 include only those that match the same set in 2003/2004. Assets in 2000/2001 should not be compared to other years since it is a reduced set of assets measured.

Table 23: Partial household-Level and National Maize Fertiliser Information:

Type of Seller	Households			Households Producing Maize		Households Producing Maize With Fertiliser	
	Ter-cile	% of HHs	Number of HHs	Number of HHs	% of HHs	Number of HHs	% of HHs
1. Grower and Seller of Maize*	Low	4.2	61,977	61,977	100	25,159	40.6
	Med	8.6	128,450	128,450	100	65,755	51.2
	High	14.7	218,885	218,885	100	158,366	72.4
Sub Total		27.4	409,313	409,313	100	249,281	60.9
2. Grower and Buyer of Maize or Mealies **	Low	14	208,786	208,786	100	37,764	18.1
	Med	13.6	202,868	202,868	100	44,043	21.7
	High	11.3	169,507	169,507	100	66,642	39.3
Sub Total		38.9	581,160	581,160	100	148,449	25.5

Source: Supplemental Survey to the 1999/2000 Post Harvest Survey, Central Statistical Office, 2007/2008 Marketing Season. Productive assets include only those that match 2004.

Table 24: Household-Level and National Maize Fertiliser Information: National Household-Level Net Yearly Income and Related Food Security

Seller type	Households			Households Producing Maize		Households Producing Maize With Fertiliser	
	Ter-cile	% of HHs	Number of HHs	Number of HHs	% of HHs	Number of HHs	% of HHs
1. Grower and Seller of Maize*	Low	3.2	39,416	39,416	100	15,364	39
	Med	8.7	108,515	108,515	100	40,345	37.2
	High	13.7	169,830	169,830	100	103,904	61.2
Sub Total		25.5	317,761	317,761	100	159,613	50.2
**2. Grower and Buyer of Maize or Mealies ** **	Low	11.1	137,680	137,680	100	27,688	20.1
	Med	8.8	109,801	109,801	100	25,342	23.1
	High	8.8	109,656	109,656	100	50,589	46.1
Sub Total		28.7	357,137	357,137	100	103,619	29

Source: Supplemental Survey to the 1999/2000 Post Harvest Survey, Central Statistical Office, 2003/2004 Marketing Season

Table 25: Partial Household-Level and National Maize Fertiliser Information: National Household-Level Net Yearly Income and Related Food Security Categorization Indicators for Zambian Rural Cropping Households According to Their Position in MAIZE AND MEALIES Market Cat

Type of Maize Seller	Households			Households Producing Maize		Households Producing Maize With Fertiliser	
	Ter-cile	% of HHs	Number of HHs	Number of HHs	% of HHs	Number of HHs	% of HHs
1. Grower and Seller of Maize*	Low	4	43,983	37,950	86.3	7,577	17.2
	Med	8.7	96,589	91,327	94.6	22,195	23
	High	12.9	143,699	141,073	98.2	72,709	50.6
Sub Total		25.6	284,271	270,351	95.1	102,481	36.1
2. Grower and Buyer of Maize or Mealies **	Low	9.1	101,528	93,907	92.5	14,458	14.2
	Med	9	100,369	96,897	96.5	19,219	19.1
	High	10	111,443	106,457	95.5	37,997	34.1
Sub Total		28.2	313,340	297,261	94.9	71,674	22.9

Source: Supplemental Survey to the 1999/2000 Post Harvest Survey, Central Statistical Office, 2003/2004 Marketing Season.

Table 26: Statistics on Re-interview Rates

		2001, Household match with	2004, Household match with	2008, Household match with
N	Valid	6922	5419	4570
	Missing	0	1503	2352

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 27: 2004, Household match with 2001 households

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	hh not found	1503	21.7	21.7	21.7
	0 matches	5342	77.2	77.2	98.9
	does not match 2001	77	1.1	1.1	100.0
	Total	6922	100.0	100.0	

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 28: 2008, Household match with 2001 households

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	hh not found	2352	34.0	34.0	34.0
	matches	4506	65.1	65.1	99.1
	does not match 2001	62	.9	.9	100.0
	2 matches 2001 but not 2004	1	.0	.0	100.0
	3 no match 2001 & 2004	1	.0	.0	100.0
	Total	6922	100.0	100.0	

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 29: Cross-tabulation, Percentile Group of area * 2008, Household match with 2001

		2008, Household match with 2001					
		hh not found	matches	does not match 2001	matches 2001 but not 2004	no match 2001 & 2004	Total
Percentile Group of area	1	557	728	9	0	0	1294
	2	531	943	12	0	0	1486
	3	449	856	10	0	0	1315
	4	428	926	18	1	0	1373
	5	339	1011	13	0	1	1364
Total		2304	4464	62	1	1	6832

Source: 2001, 2004, 2008 Supplemental Survey FSRP/CSO/MACO

Table 30: Cross-tabulation year * Was Govt basal dressing fert available on time?

			Was Govt basal available on time?		
			No	yes	Total
year	1999	Count	204	173	377
		% within year	54.1%	45.9%	100.0 %
	2002	Count	94	228	322
		% within year	29.2%	70.8%	100.0 %
		l			
	2006	Count	258	707	965
		% within year	26.7%	73.3%	100.0 %
Total		Count	556	1108	1664
		% within year	33.4%	66.6%	100.0 %

Table 31: Cross-tabulation year * Was Govt top dressing fert available on time?

			Was Govt top available on time?		
			No	yes	Total
year	1999	Count	204	173	377
		% within year	54.1%	45.9%	100.0 %
	2002	Count	119	205	324
		% within year	36.7%	63.3%	100.0 %
	2006	Count	248	725	973
		% within year	25.5%	74.5%	100.0 %
Total		Count	571	1103	1674
		% within year	34.1%	65.9%	100.0 %
		% of Total	34.1%	65.9%	100.0 %

Figure 8: Area Planted to Maize in the 1999/2000 season by Small and Medium Scale Households

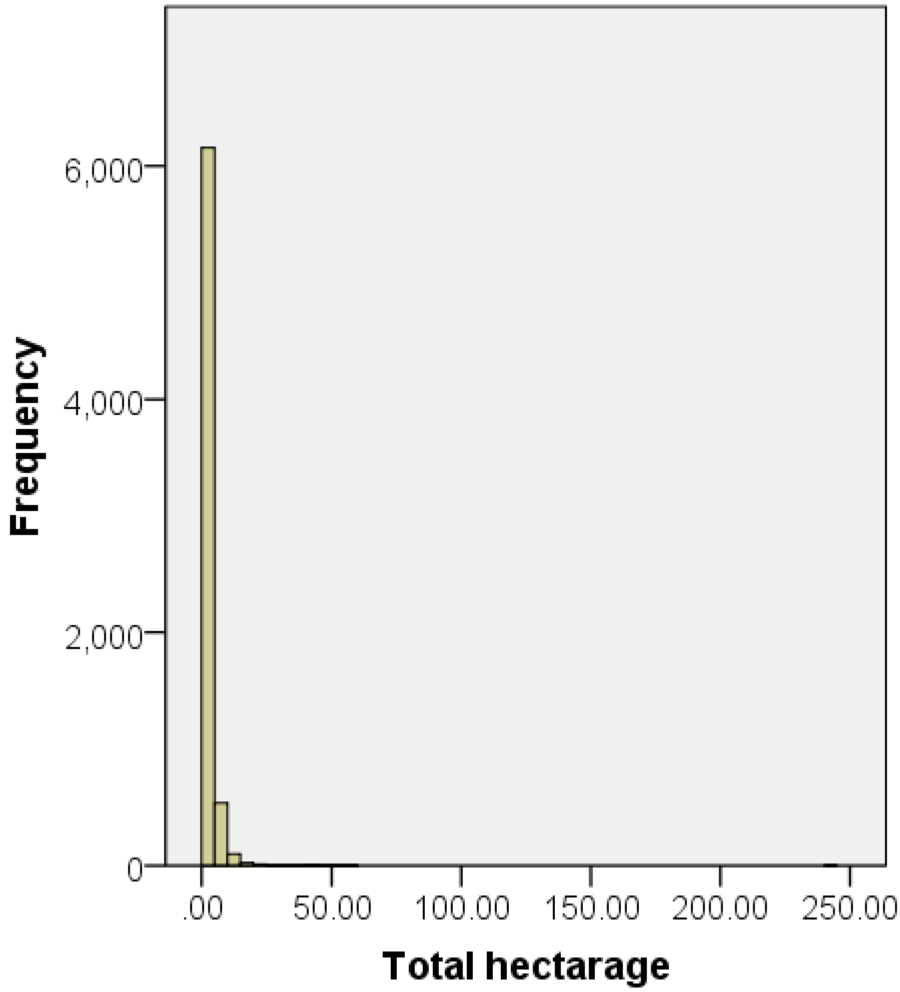


Figure 9: Area Planted to Maize by Small & Medium Scale Households in the 2002/2003 Season

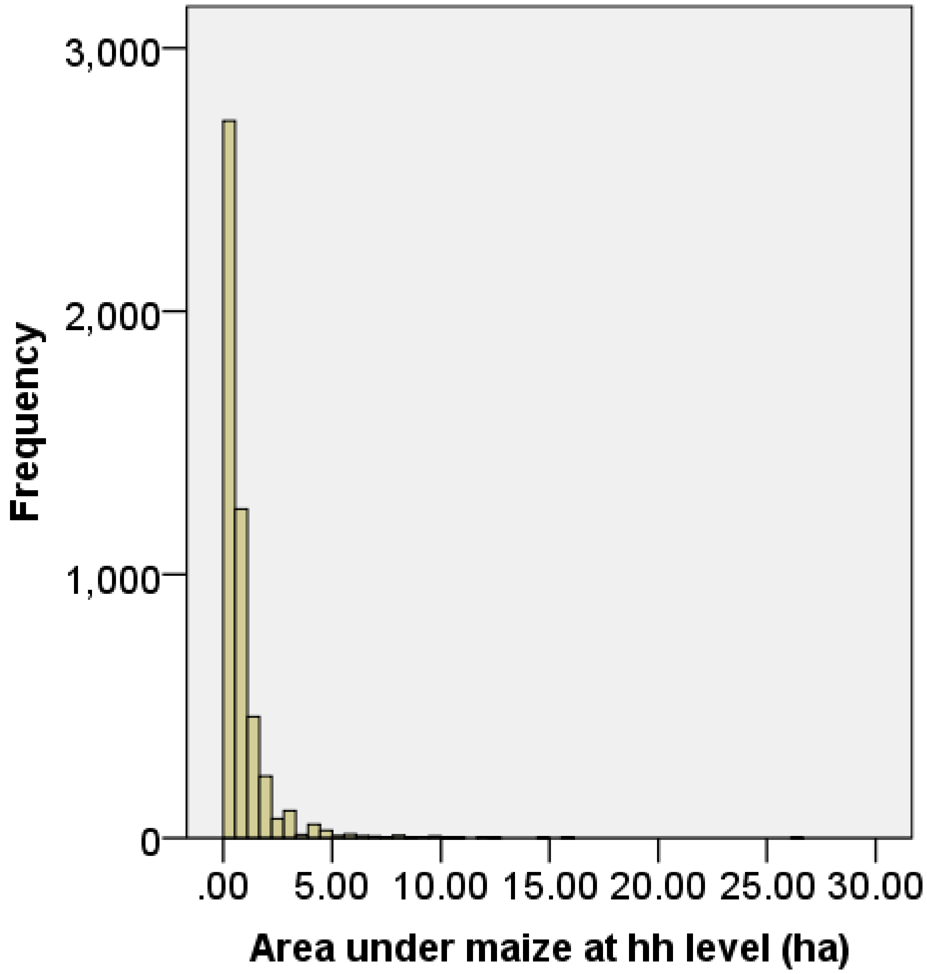


Figure 10: Maize Yield based on 2002/2003 Season

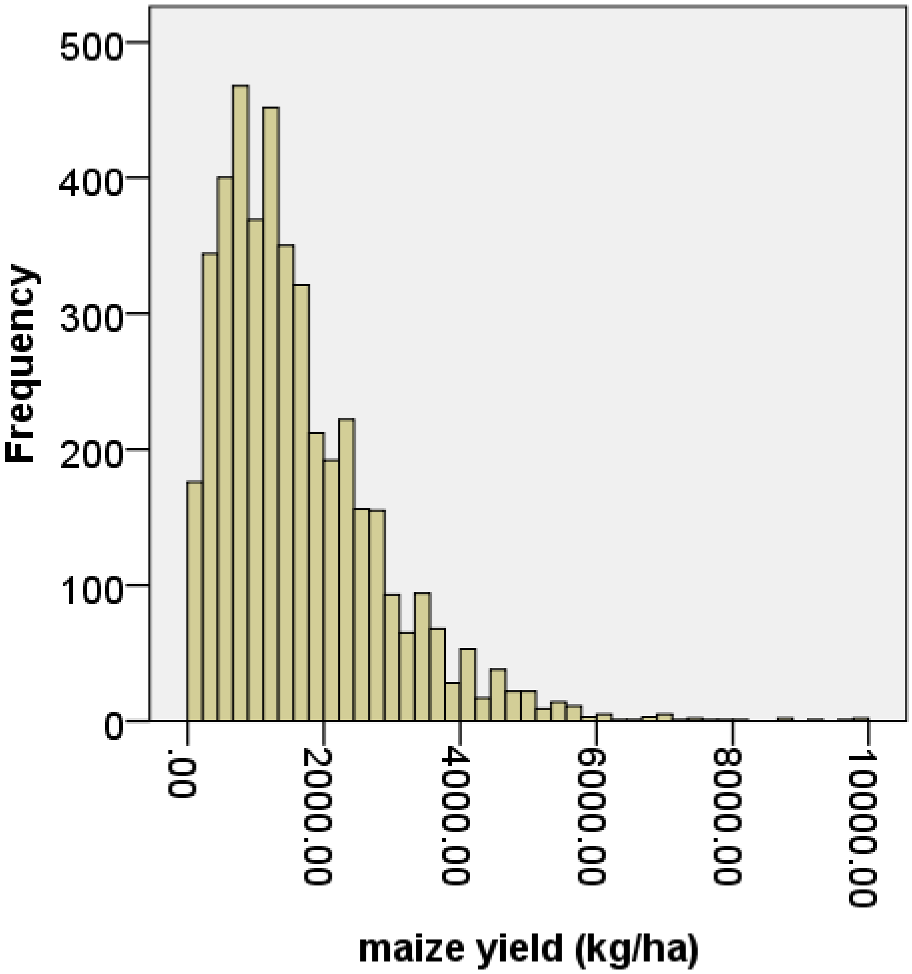


Figure 11: Maize Yield per ton based on 2006/2007 season

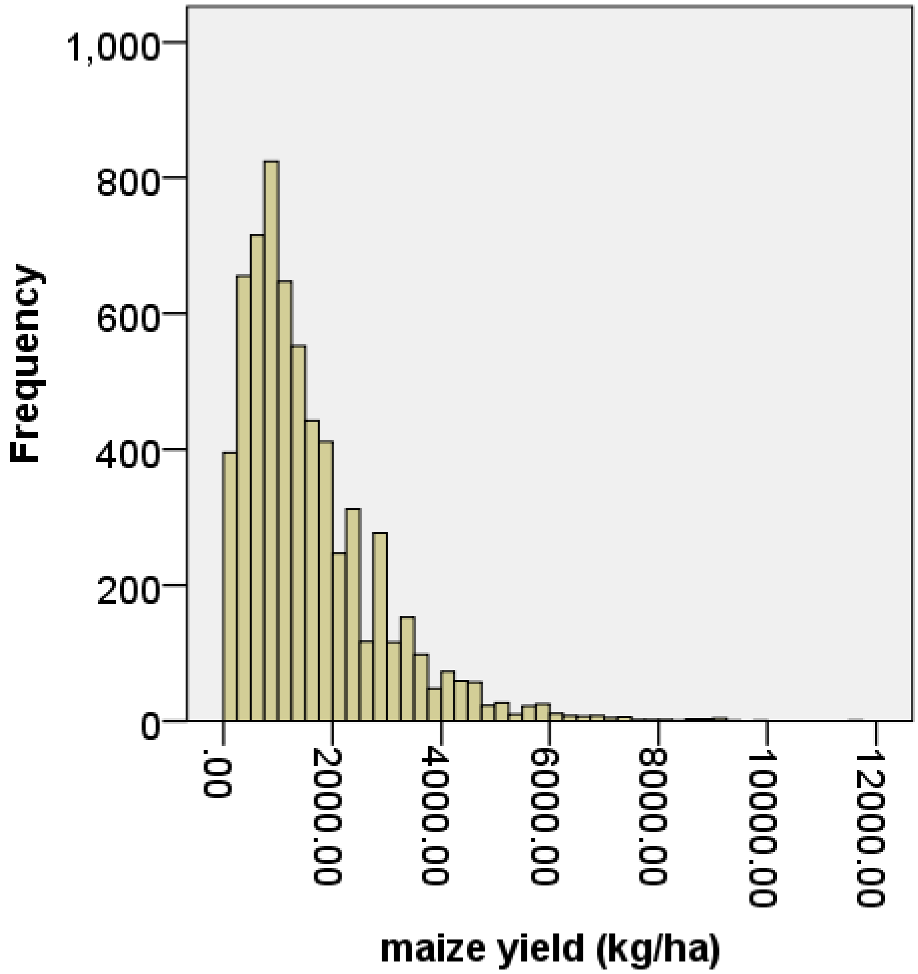


Figure 12: Total Household Maize output based on 1999/2000 season

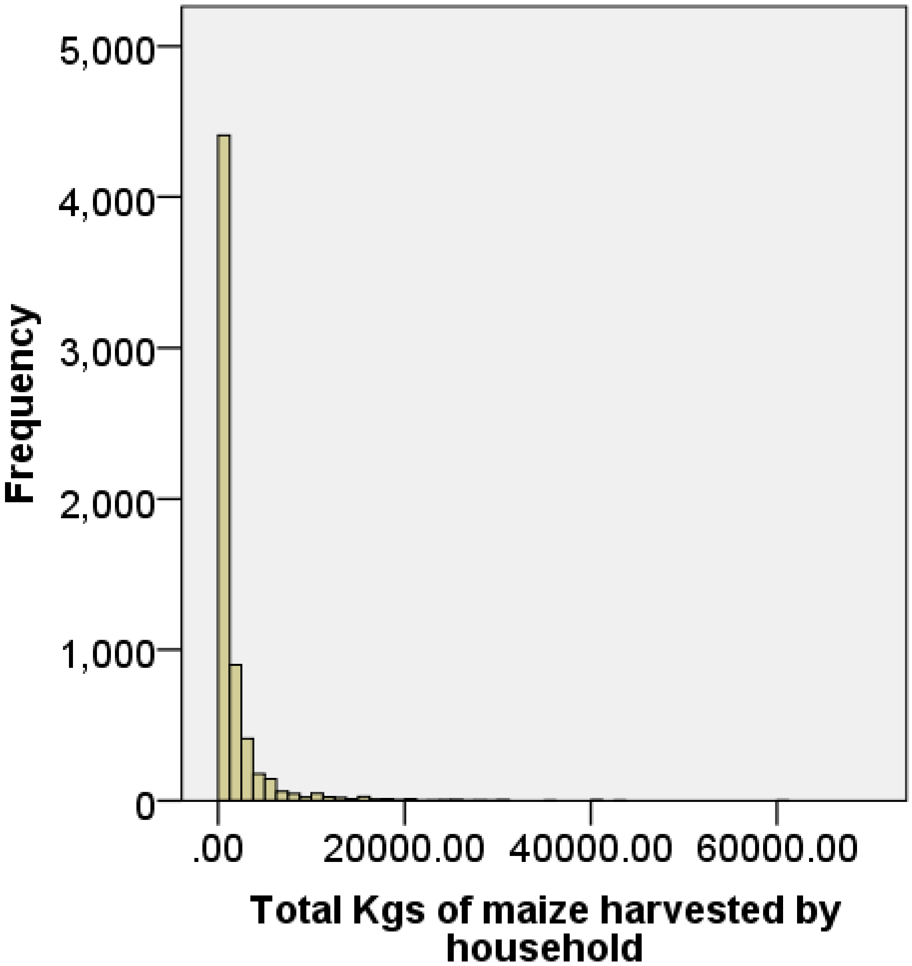


Figure 13: Total Maize output based on 2002/2003 season

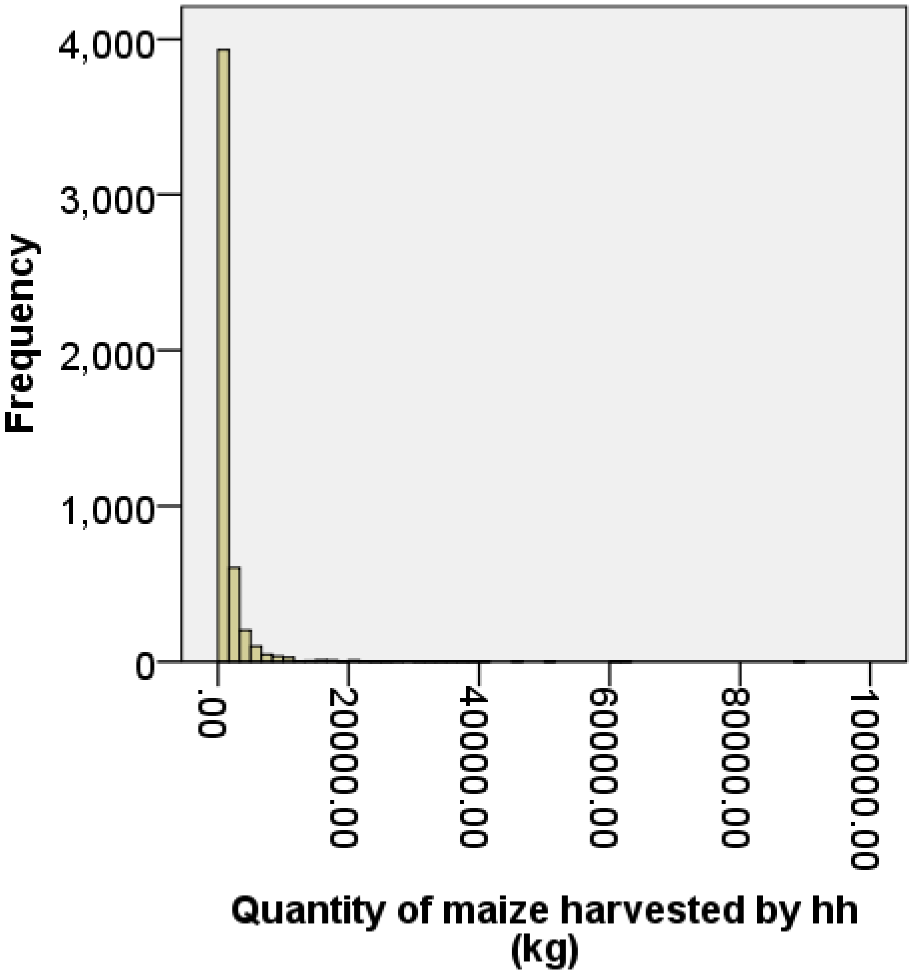
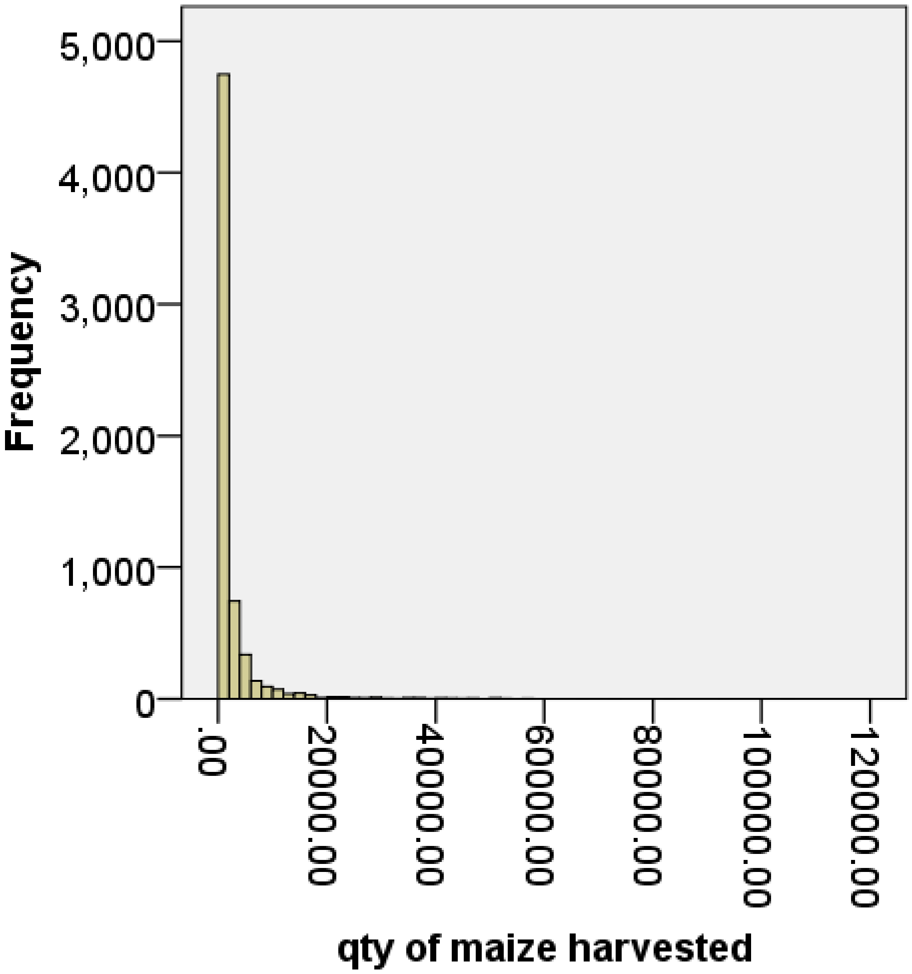


Figure 14: Total Maize harvest based on 2008 supplemental survey season



APPENDIX B: SAMPLE SELECTION

The sample for the supplemental surveys was drawn from all current 72 districts of Zambia. The country is divided into the following administrative units, province, district and ward. For the purpose of sampling, the Central Statistical Office (CSO) has further subdivided the wards into Census Supervisory Areas (CSA) and Standard Enumeration Areas (SEA). 'The SEA is the smallest area with well-defined boundaries identified on census sketch maps'. [iii]An SEA contains approximately between 100 -150 households and 20 households are sampled from each SEA. A stratified multi-stage sample design was used for the Zambia PHS. 'The sampling frame was based on the data and cartography from the 1990 Census of Population, Housing and Agriculture. The primary sampling units (PSUs) were defined as the Census Supervisory Areas (CSAs) delineated for the census.

The CSAs were stratified by district within province and ordered geographically within district. A total sample of 405 CSAs was allocated to each province and district proportionally to its size (in terms of households). A master sample of CSAs was selected systematically with probability proportional to size (PPS) within each district at the first sampling stage; the measure of size for each PSU was based on the number of households listed in the 1990 Census. The secondary sampling unit is the Standard Enumeration Area (SEA), defined as the segment covered by one enumerator during the census. One SEA was selected within each sample CSA with PPS for the survey'.[iv] An average SEA contains between 150 – 200 households. Once an SEA was selected, an enumerator visited all the households within the SEA and collected basic information about the total area cultivated by the household.

SELECTION OF SAMPLE HOUSEHOLDS

The first step is to identify agricultural households among those listed in the SEA, i.e. households that reported having grown crops, and /or raised livestock, and/or raised chickens. Households that are non-agricultural, those that are 'non-contact' and those that refused to cooperate should also be identified and indicated by writing 'NON AGRIC' 'NON CONTACT' or 'REFUSAL' in the margin against them. Put a mark in the relevant column under column 11 for households that have been identified as either 'NON AGRIC' 'NON CONTACT' or 'REFUSAL'.

The next step is to stratify agricultural households by size of cultivated land (column 7) and, in certain cases, on the growing of some specified crops (column 8), on numbers of cattle, pigs and goats raised (column 9) and on number of chickens raised (column 10). The agricultural households will be stratified into three (3) categories: A, B and C.

Category C: Area under crops 5.0 – 19.99 ha

This category will also includes households reporting any of the specified crops when only 1 or 2 households in the SEA report the specified crop(s), even if they do not qualify basing on area under crops. Households raising 50 or more cattle, 20 or more pigs, 30 or more goats, and/or 50 or more chickens, even if they do not qualify basing on area under crops.

Category B: Area under crops 2.0- 4.99 ha

This category will also include households reporting any of the specified crops, when 3 to 5 households in the SEA report the specified crop(s), even if they do not qualify basing on area under crops.

Category A: All the remaining agricultural households with area under crops less than 2.0 hectares.

Stratification Procedure

When stratifying households, start with category C.

Identify the households that reported area under crops (column 7) of 5.0 to 19.99 hectares and put a mark (x) in category C column under column 11 in the row of each of such households. Identify the households that reported any of the specified crops (column 8). Count such households. If there are only 1 or 2 such households, include them in Category C by putting a mark (x) in category C column under column 11 in the row of these households.

Using column 9, identify households that reported raising 50 or more cattle, 20 or more pigs, 30 or more goats and treat these in the same manner as explained in '2'.

Using column 10, identify households that reported raising 50 or more chickens and treat these in the same manner as explained in '2'.

Category B

Identify households that reported area under crops (column 7) of 2.0 to 4.99 hectares and put a mark (x) in category B column under column 11 in the row of each such households.

Identify households that reported any of the specified crops (column 8). Count such households. If there are 3 to 5 such households, include them in Category B and put a mark (x) in category B column under column 11 in the row of each of these households.

NOTE: if there are more than 5 households in an SEA reporting any of the specified crops, these households will not automatically be included in category 'C' OR 'B' but stratification will be based only on area under crops.

Category A

First critically check the stratification of households in category C and B and when you are satisfied that everything is in order, all the remaining households have area under crops of less than 2 hectares, are among the more than 5 households reporting any of the specified crops, and have reported raising less than 50 cattle, less than 20 pigs, less than 30 goats and less than 50 chickens. All such households belong to category A. Put a mark (x) in category A column under column 11 in the row of each of these households.

Assign Sampling Serial Numbers, within each category, following where you put (x). The sampling serial numbers will sequentially be assigned, starting with '1' in each category. In addition assign serial numbers to 'NON AGRIC ' households in the appropriate column in col. 11 and then do the same for 'NON CONTACT' and 'REFUSAL' households in the 'NON CONTACT' column.

NOTE: (a) The sum of the last sampling serial numbers in categories A,B and C must be equal to the total number of agricultural households listed in the SEA.

(b) The sum of the last serial numbers in col. 11 must be equal to the last household serial number in the SEA.

Summary of Households Listed in SEA

Column 1. Gives the categories as allocated to households in Col. 11 of the Listing Book. Note that 'Non-Contact includes refusals.

Column 2. Enter, against each category, the serial number assigned to the last household in the category (Col.11).

Enter the sum of categories A,B,C and 'Non Agric' against 'SUB-TOTAL'. This will give the number of households that gave complete response during listing.

Add 'Non Contact' to 'Sub-Total' and enter the result against 'TOTAL'. This gives the total number of households in the Sea i.e it should be equal to the serial number assigned to the last household listed.

Columns 3,4,5. Completing of these columns is explained under 'Sample households Selection'. Sample households will be selected from categories A,B and C under Col. 11 of the Listing Book. This means that the sample will be drawn only from agricultural households that gave complete response during the listing exercise.

Sample Household Selection

The total sample size in each SEA is 20 households. Where all the three categories have adequate numbers of households (10 or more) listed, the sample households distribution will be, C–10, B–5 and A– 5.

In cases where there are shortfalls in category C, include all households in this category and allocate the difference from 20 equally to categories B and A. if the differences from 20 cannot be equally allocated to the two categories, allocate category B one (1) more sample household than category A

Where there is no household in category C, allocate 10 sample households to category B, and 10 to category A.

Where there is no household in category C and less than 10 in category B, include in the sample all those in B and increase the allocation in category A to make up for the shortfall from the required number of 20 sample households.

Where all households in an SEA fall in category A, select all the required 20 sample households from that category

Systematic Sampling Procedure

The allocated number of sample households to each category will be selected independently using the following procedure:

Divide the total number of households listed in the category by the number of households to be selected (according to sample allocation) to give the Sampling Interval (SI). Calculate this to two (2) decimal places. From the table of random numbers, get a random number (RS) between ‘1’ and the SI, inclusive. The random number obtained will give the first household that will be in the sample. Add the SI to the random number (RS), and the integer part of the sum will give the second household to be in the sample.

Continue with the procedure, adding SI to each successive sum until you have all the allocated sample size for the category. Put a circle round each sampling serial number (column 11), in the listing book, corresponding to the numbers you have worked out for each category. The sampling serial numbers circled will indicate the households selected for the sample.

Transcribe onto the ‘LIST OF SELECTED HOUSEHOLDS’ sheet, now copying the household serial numbers (column 2) of the selected households.

REFERENCES

REFERENCES

Anthony Chapoto, T.S.J., N. Mason (2007). Security of Widows' Access to Land in the Era of HIV/AIDS: Panel Survey Evidence from Zambia (Revised Version). Food Security Research Project, pp 4.

Brooks, J., G. Dyer and E. Taylor (2008). Modelling Agricultural Trade and Policy Impacts in Less Developed Countries, *OECD Food, Agriculture and Fisheries Working Papers*.

Chavas, J.-P. (1995). *The Microeconomics of Food Security*. Mimeo, Department of Agricultural Economics University of Wisconsin-Madison.

Chavas, L.C.S.a.J.-P. (1999). SUPPLY RESPONSE OF WEST AFRICAN AGRICULTURAL HOUSEHOLDS: IMPLICATIONS OF INTRAHOUSEHOLD PREFERENCE HETEROGENEITY, (Serial online)FCND DISCUSSION PAPER NO. 69). Available from URL.

Chitalu, Z. (2000). *Databases*.

Deaton, A. (1997). *The Analysis of Household Surveys*. The Johns Hopkins University Press.

DeJanvry, F.S. Market Failure.

Zambia Meteorological Department

Economist, T. (April 17th 2008). Hand to mouth, *The Economist Magazine*.

Economist, T. (April 17th 2008). The new face of hunger, *The Economist*.

Ehui, S., S. Benin, and Z. Paulos. (2003). Policy Options for Improving Market Participation and Sales of Smallholder Livestock Producers: A Case Study of Ethiopia. Paper Submitted to the 2nd EAF International Conference on Contemporary Development Issues in Ethiopia, Addis Ababa, Ethiopia, 11-13 July.

Filmer, D.P., L. H. (1998). The effect of household wealth on educational attainment: Demographic and health survey evidence. World Bank.

Goetz, S.J. (1992)). A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa, *American Journal of Agricultural Economics* 74: 444-452.

Hallam, D. Analysing Agricultural Supply Response In Economies In Transition, *Options Méditerranéennes* Serie B: 123-131.

Holloway, G.J., C. Nicholson, C. Delgado, S. Staal, and S. Ehui. (2000). Agroindustrialisation through Institutional Innovation: Transaction Costs, Cooperatives and Milk-Market Development in the East African Highlands, *Agricultural Economics* 23: 279-288.

Inderjit Singh, L.S.a.J.S. (1986). *Agricultural Household Models -Extensions, Applications and Policy*. The Johns Hopkins Press.

Irma A. Gomez, H.A.L.a.D.M.B. (1999). Alternative Price Expectation Regimes in Timber Markets, *Journal of Forest Economics* 5.

J. Ricker-Gilbert, T.S.J., E. Chirwa (2009). How Do Fertilizer Subsidies Affect the Demand for Commercial Fertilizer? An Example from Malawi.

Jayne (2007). Smallholder Household Maize Production And Marketing Behavior In Zambia: Implications For Policy.

Jayne, G.T.T.S. (2009). Agricultural Trade Flows among developing countries: Do Regional Preferential Trade Agreements make a Difference?

Langyintuo, A.S. (2008). Computing Household Wealth Indices Using Principal Components Analysis Method. International Maize and Wheat Improvement Centre (CIMMYT), Harare, pp 3.

Larson, Y.M.a.D.F. (1992). On the Transmission of World Agricultural Prices, *The World Bank Economic Review* Vol. 6: 399-422.

Liapis, P. (2009). Extensive Margins in Agriculture. in OECD Food, A.a.F.W.P. (ed.). OECD.

Mather, Boughton, Jayne (Smallholder Heterogeneity and Maize Market Participation in southern and Eastern Africa: Implications for Investment Strategies to Increase Marketed Food Staple Supply, MSU International Development Working Paper 113

Martin, M.I.W. (2008). Implications of Higher Global Food Prices for Poverty in Low-Income Countries. The World Bank Development Research Group.

McFadden, D. (1999). Instrumental Variables. UCLA Berkeley.

Megill, D. Recommendations On Sample Design For Post-Harvest Surveys In Zambia Based On The 2000 Census. Food Security Research Project.

Nicholson, W. (2005). *Micro-economic theory Basic principles and Extensions*.

Nicky Mason, T.J., Mike Weber, James Shaffer (2008). Patterns of Maize Farming Behaviour and Performance Among Small and Medium-Scale Smallholders in Zambia, *Food Security Research Project, Zambia*, pp 14.

OECD (2008). Rising Food Prices: Causes and Consequences.

Pearl, J. (2000). Causality: Models, Reasoning, and Inference, *Cambridge University Press*.

Persaud, W.L.S. (2009). The Transmission of Exchange Rate Changes to Agricultural Prices. United States Department of Agriculture.

Pritchett, D.F.a.L.H. (2001). Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India, *Demography* 38: 115-132.

Food Security Research Project, (2009). Zambia Maize Value Chain Study.

RAHJI M. A. Y and Adewumi, M.O. (2008). Market Supply Response And Demand For Local Rice In Nigeria: Implications For Self-Sufficiency Policy, *Journal of Central European Agriculture* 9: 567-574.

Sadoulet, A.d.J.a.E. Progress in the Modelling Of Rural Households' Behaviour Under Market Failures.

Snijders, T.A.B. (2005). Fixed and Random Effects. in B.S. Everitt and D.C. Howell (eds.) (ed.), *Encyclopedia of Statistics in Behavioral Science.*, pp 664-665.

Squire, H.N.B.L. (1980). Predicting Agricultural Output Response, *Oxford Economic Papers,* 32: 284-295.

USDA (2007). USDA Agricultural Projections to 2016. in Economist, O.o.t.C. (ed.). USDA.

USGS (2002). Water Requirement Satisfaction Index, (Serial online). Available from URL: earlywarning.usgs.gov/centralamerica/readme/WRSI_readme.doc [accessed.

Whitehead, A. (1990). *Food Crisis and Gender Conflict in the African Countryside*. Monthly Review Press, New York.

Whitehead, A. (1990). *Rural Women and Food Production in Sub-Saharan Africa*. Clardendon Press, Oxford.

Williams, R. Notre Dame University

Wooldridge, J.M. (2005). *Introductory Econometrics*. South West College Publishing.