# AN ESTIMATION OF SUGARCANE SUPPLY RESPONSE AMONG OUTGROWERS IN MUMIAS SUGAR COMPANY IN KENYA

**MUTUA MERCY MULI**

**A thesis submitted in the Department of Agricultural Economics, Faculty of Agriculture, in partial fulfillment of the requirements for the award of Master of Science degree in Agricultural and Applied Economics, University of Nairobi**

**NOVEMBER 2015**

## DECLARATION

This thesis is my original work and has not been presented for a degree in any University.

Signed …………………………………………… Date ……………………

Mercy M. Mutua (BSc Agribusiness Management) A56/64447/2010

This thesis has been submitted with our approval as University Supervisors:

Dr. Patrick Irungu …....................................... .......................................

Supervisor Signature Date

Prof. Stephen G. Mbogoh ……..................................... .......................................

Supervisor Signature Date

# DEDICATION

This thesis is dedicated to my niece Shakina.

## ACKNOWLEDGEMENT

This study would not have been completed without the support of different individuals. Firstly, I am greatly indebted to my supervisors Dr. Patrick Irungu and Professor Stephen Mbogoh for their tireless effort and advice from the initial stages of proposal writing, through data collection and to the final stages of writing this thesis. Your careful scrutiny of this work was invaluable.

I also wish to acknowledge the AERC for the partial scholarship towards the third semester of my coursework at the shared facility at University of Pretoria, South Africa and for research funds which enabled me to collect the data used in this study. The resources support played a major role towards the successful completion of my MSc. studies and for that I am very grateful. Special thanks go to Mr. Mugo and Mr. Mulehi of the Kenya Sugar Board both of whom were very resourceful during the data collection.

My heartfelt gratitude goes to the Director- Tegemeo Institute, Dr. Mary Mathenge for her support in terms of time, advice and encouragement throughout the coursework and thesis writing. I'm indebted for all the support offered by the Institute. Dr. Eliud Birachi, my immediate supervisor at CIAT, was extremely understanding and supportive during my thesis defense and final submission.

My family, classmates and colleagues were a great source of encouragement and support. May God bless and reward you. Finally, I thank God for His grace and favour throughout my MSc Programme. To your name be the glory.

# ABSTRACT

Over the past few decades, Kenya has had a structural deficit in sugar despite the great potential the country has in sugarcane production. Kenya's sugar industry relies entirely on sugarcane production. Sugarcane production presents an opportunity for income generation and poverty alleviation to small scale sugarcane farmers in the cane growing areas. While the country has good agro-climatic conditions for sugarcane production, the question as to why the milling companies face shortages in sugarcane supply and often operate below their installed capacity and resort to cane poaching remains unanswered. The objective of this study was to assess the responsiveness of farmers supplying sugarcane to Mumias Sugar Company to changes in the producer environment. The study used time series data collected from the annual publications by the Kenya Sugar Board and the Ministry of Agriculture. A Vector Error Correction Model (VECM) was used to estimate the responsiveness of sugarcane farmers to prices and other economic incentives for the period 1980 to 2013.

The unit root tests showed that all the variables were integrated of order one i.e, I(1). The cointegration tests showed that the variables were co-integrated with one cointegrating equation. The presence of integration in the individual variables and cointegration among variables set the condition for the use of a Vector Error Correction Model in supply response analysis. The results showed that in the short-run, planned supply expressed in terms of land area allocated to sugarcane was positively affected by changes in its own price in the immediate preceding period. The responsiveness of the farmers to changes in own-price was however inelastic implying that manipulating the price of sugarcane may not yield much response in terms of land allocation in the short-run.

In the long-run, all variables (own price, maize price, yield, privatization dummy and time trend) in the model were found to significantly affect the area of land allocated to sugarcane

by outgrowers. The own-price and yield elasticities were positive while the cross-price elasticity of sugarcane supply was negative as hypothesized. The coefficient of the dummy variable for privatization and time trend were also negative but significant. Sugarcane supply was however found to be inelastic to all variables in the model in the long-run. This implies that the sugarcane farmers might be facing a captive value chain whereby they are dependent on the Sugar company when making land allocation decisions.

With regard to the cross-price elasticity, the maize price elasticity of sugarcane supply was less than unity. While a price reduction strategy for the competing enterprise cannot be encouraged as it would be counter productive, policies that would improve the profitability of sugarcane relative to other enterprises are recommended in order to make it more attractive to the farmers in support of government's sugar self-sufficiency policy goal. Such policies may involve investing in yield-enhancing interventions such as making inputs accessible at more competitive prices and taking control of the unfavourable impacts of weather-related risks through irrigation as well as research and development of high-yielding sugarcane varieties with high sucrose content.

A dummy variable based on the change of milling company ownership and management from government to private ownership for Mumias outgrowers was significant but negative in the long-run with respect to the farmers' land allocation decisions. This implies that in order for the policy shift of privatization to achieve increased land allocation to sugarcane, it is necessary that other impediments to the farmers' supply response be addressed. Such impediments include mismanagement of the milling company and inefficiencies in processing payments to the farmers for cane delivered. It is recommended that a study similar to this one is done after some years when a longer time series on the privatization variable is available to ascertain whether the status remains.

The long-run elasticity of supply estimates were found to be higher than the short-run estimates. This was expected because land is a fixed asset and its reallocation moves towards the desired level as factors that are fixed in the short-run become more variable in the long-run. Based on the findings of this study, it can be concluded that the sugarcane farmers supplying Mumias Sugar Company are facing a captive value chain. Their dependence on the company limits their responsiveness to changes in the producer environment such that while they respond as expected, all the elasticities are inelastic. It may be time to relax the self-sufficency policy goal in sugar and encourage farmers to grow what they consider profitable so as to put an end to the welfare losses emanating from the sub-sector.

## TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

# LIST OF ACROMYMS

| | |
|---|---|
| ADF | Augmented Dickey Fuller |
| ADL | Autoregressive Distributed Lag Model |
| AERC | African Economic Research Consortium |
| CET | Common External Tariff |
| CMAAE | Collaborative Masters in Agriculture and Applied Economics |
| COMESA | Common Market for Eastern and Southern Africa |
| CPI | Consumer Price Index |
| EAC | East African Community |
| ECM | Error Correction Mechanism |
| GAIN | Global Agricultural Information Network |
| GDP | Gross Domestic Product |
| DFGLS | Dickey-Fuller Generalized Least Squares |
| GIS | Geographical Information Systems |
| KSB | Kenya Sugar Board |
| ML | Maximum Likelihood |
| NCPB | National Cereals and Produce Board |
| OLS | Ordinary Least Squares |
| RMSE | Root Mean Square Error |
| SDL | Sugar Development Levy |
| SONY | South Nyanza Sugar Company |
| USDA | United States Department of Agriculture |
| VAR | Vector Autoregression |
| VAT | Value Added Tax |
| VECM | Vector Error Correction Model |

# CHAPTER 1: INTRODUCTION

## 1.1 Background

### 1.1.1 Sugarcane production in Kenya

Kenya's sugar industry relies heavily on domestic sugarcane production. About 92 percent of sugarcane in Kenya is produced by smallholder farmers, with the rest of production being undertaken by a combination of large scale farmers and nucleus estates that are owned by milling companies (Kenya Sugar Board, 2011). The industry is composed of five parastatal companies which include Nzoia, South Nyanza (SONY), Chemelil, Muhoroni[1] and Miwani sugar companies and eight privately run sugar companies, i.e., Mumias[2], Kibos Sugar & Allied Industries Ltd, West Kenya, Soin, Butali, Kwale International Sugar Company, Transmara Sugar Company and the newly established Sukari Industries Ltd.

Figure 1.1 shows the area under sugarcane, production and yield trends for the period between 1992 and 2013. Since 2000, the total national sugarcane production and the area harvested have been increasing. This is mainly as a result of an increase in total land under sugarcane cultivation (FAO, 2013). Data from FAOSTAT and the Kenya Sugar Board (KSB), however, show that despite the increase in production, the productivity started declining in 2008 (Figure 1.1). This scenario implies that the increase in output from 2008 was mainly as result of increased land under cultivation as opposed to yield gains. Odenya *et al*. (2010) cites the low adoption of improved cane varieties as one of the main causes of the decline in productivity. Other reasons for the declining yields include low use of fertilizers and certified seed cane owing to high prices, poor agricultural and land management practices and delayed harvesting of mature sugarcane (KSB, 2010).

[1] Muhoroni amd Miwani Sugar Companies are in receivership
[2] The government's stake in Mumias Sugar Company was reduced from majority shareholding to only 20 percent when the company was privatized in 2001 (Mumias Sugar Co. website).

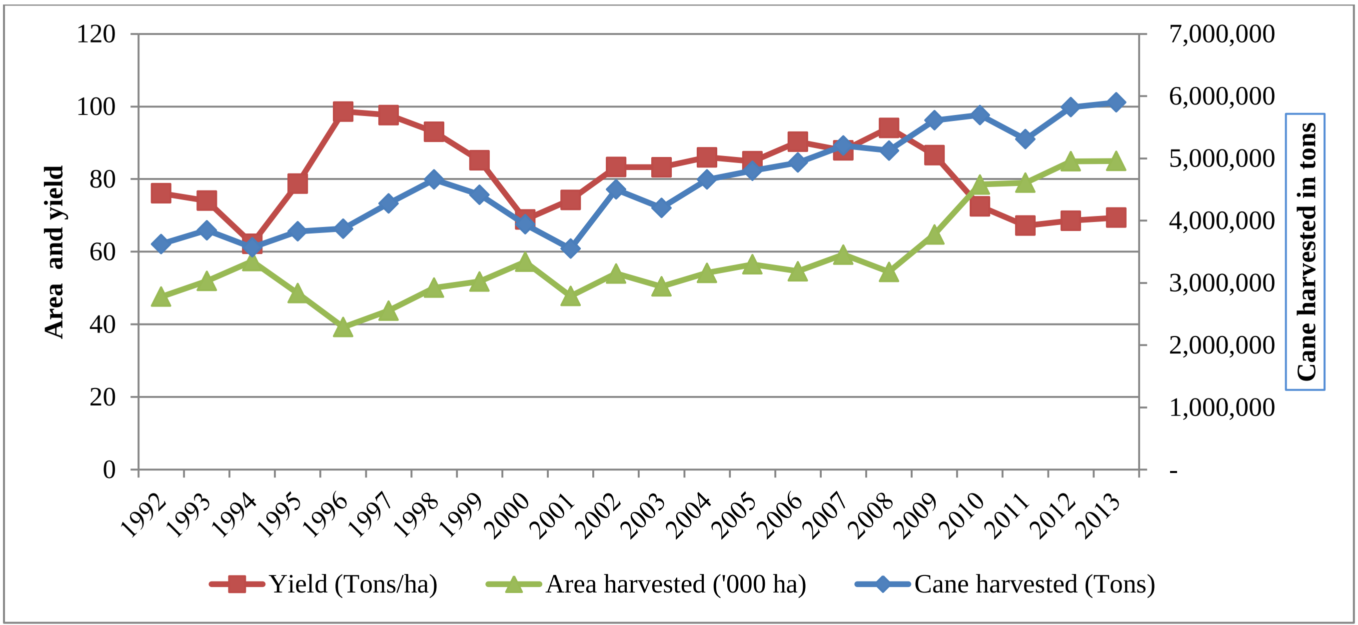



**Figure 1.1: Area under sugarcane production and productivity in Kenya (1992-2013)**
Source: FAOSTAT and KSB (2013)

Kenya's sugar production has for a long time fallen short of domestic demand with the gap, averaging 200,000 metric tons (MT) per year, being bridged through imports. A report on the Kenya sugar industry by the Global Agricultural Information Network (GAIN) notes that local production only supplies up to 70 percent of domestic consumption. The country has the highest per capita consumption of sugar in the East African Community (EAC) at 21 kilograms per person per year (USDA, 2011) against an average of 13 and 10 kilograms per capita in Tanzania and Uganda respectively.

Figure 1.2 shows trends in sugar production, consumption, exports and imports for the period 1995 to 2012. The consumption has been consistently higher than domestic production, with the deficit being met through imports. The consumption has been on the increase owing to growth in population, income and urbanization as well as growth in industrial and food service sectors (USDA, 2014). The upward trend in sugar production is a result of increasing sugarcane production as observed in Figure 1.1.

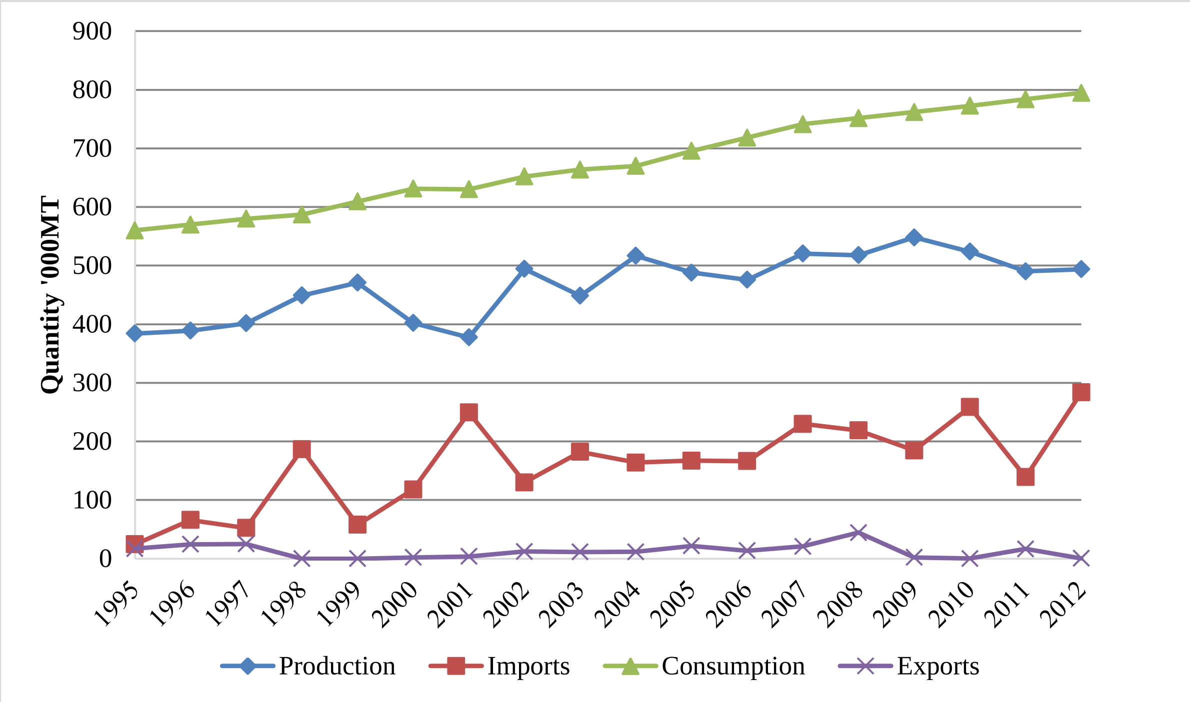



**Figure 1.2: Trends of sugar production, consumption, imports and exports in Kenya (1995-2012)**
Source: KNBS (Various issues)

### 1.1.2 Policy issues in Kenya's sugar sub-sector

The sugarcane sub-sector contributes about 15 percent of the agricultural gross domestic product (AgGDP) and supports approximately 250,000 small scale farmers in terms of livelihood (KSB, 2011). About six million people are directly employed in the sugar value chain, from production to consumption (KSB, 2011). The sub-sector is an import substituting industry and hence it saves the country millions of shillings annually (KSB, 2004).

Despite the importance of the sugar sub-sector in Kenya's economy, it is marred with challenges that have not allowed it to be competitive relative to other sugar producers in the region. These include high production costs and low capacity utilization by milling companies due to lack of sufficient raw materials. There are also cases of cane poaching, where farmers deliver cane to a company different from their contractor who had provided them with raw materials on credit. This makes it difficult for the contracting company to recover the inputs. The major reason for the low competitiveness of local sugar companies is inefficiency in both cane production (Mulwa *et al.*, 2005) and processing (Gicheru *et al.*, 2007). Accordingly, Kenya has one of the highest sugar production costs in the region (RoK, 2010), averaging $600 per MT (USDA, 2015) against an average of USD 350 per MT in other COMESA countries.

The Kenyan sugar market is protected through tariff and non-tariff barriers. The former is in the form of 100 percent *ad valorem* tariff applied on sugar imports in excess of 220,000 MT annually. Non-tariff barriers include registration and licensing of sugar importers (at a cost of KShs 100,000 a year), import permits acquired from the KSB for every consignment, Value Added Tax (VAT) of 16 percent, import quotas for raw and refined sugar, and sugar development levy (SDL) of seven percent that applies to non-industrial sugar importers (USDA, 2011).

In order to effect the tariff barriers, Kenya presented a case to the COMESA on the need to protect the sub-sugar sector since 2000. This was achieved by entering into a special agreement on sugar import safeguard with COMESA to initially import 200,000 MT of duty-free sugar yearly from COMESA member countries (Odhiambo, 2008). This quantity was sufficient to bridge the excess demand in the country up to 2007. The gap in sugar demand has however increased since 2007 to 220,000 MT annually and the import quota subsequently raised. All imports in excess of the import quota were to be taxed at the EAC common external tariff (CET) of 100 percent. The original plan was to gradually zero-rate the import tariff by 2008. This move was meant to protect the local sugar industry, which remains uncompetitive due mainly to production, management and processing related inefficiencies.

The sugar import safeguard period was however further extended from 2008 to February 2012 and since then Kenya has successfully been awarded yearly extensions by COMESA.

The import safeguard measure has, however, been found to lead to inefficiencies in the sub-sector (Kipruto, 2010). While it may generally be viewed as a positive step, especially in the interest of farmers, it cannot be assumed that the import safeguard will continue to be allowed by COMESA indefinitely. There is therefore need to come up with a more sustainable and efficient policy to steer the sub-sector to competitiveness.

The main proposition that has been put forward to steer the sugar sub-sector towards greater competitiveness in anticipation of complete liberalization of trade within the COMESA is the privatization of Kenya’s sugar factories. The Government of Kenya (GoK) intends to offload part of its shareholding in the five publicly-owned sugar companies, i.e., Nzoia, SONY, Chemelil, Muhoroni and Miwani. The privatization plan was approved by the cabinet in 2012 and in 2015 the government advertised bids for strategic investors to express their interest. In the proposed ownership arrangement, the government plans to off load 51% stake to strategic investors, 24 percent to employees and farmers and retain 25 percent (The Business Daily Newspaper).

Since the early 1960s, Kenya’s national policy on the development of the sugar industry has been based on promoting domestic self-sufficiency (Mbogoh, 1980). This goal was achieved only between 1980 and 1981 (Nyongesa, 2003). This was mainly as a result of investments made in the previous century coming to fruition with the entrance of SONY and Nzoia Sugar companies into the market (Hornsby, 2012). Thereafter, the country has largely been unable to remain self-reliant in sugar production. Between 1972 and 1992, the government used to fix sugar cane prices and ex-factory as well as consumer prices of sugar in order to protect both farmers and consumers (Mbogoh, 1980). The emphasis was to give farmers incentives to grow more cane because the millers were operating under capacity (Nyongesa, 2003). In the later 1990s, however, the policy shifted to encouraging more investment in the industry to expand processing capacity.

Sugarcane production presents an opportunity for income generation to small scale sugarcane farmers in cane growing areas. While Kenya has good agro-climatic conditions for sugarcane production, the question as to why there are frequent shortages of sugarcane supply to factories remains unanswered. In recent times, Kenya has mainly been following two policy alternatives to manage the structural deficit in sugar: the protection of the farmers through import safeguards, and the privatization of publicly owned sugar companies. Several studies have made recommendations targeting farmer behavior in order to improve sugar cane supply. For any such measures to have the intended impact on farmer behavior, it is important to understand how farmers are likely to respond to it. Muchapondwa (2009) argues that the application of agricultural policy instruments on agricultural activity without an empirical understanding of the structural parameters of supply may lead to unintended results.

A pricing policy, for instance, can be used to achieve increased production for a commodity, but for this to happen there is need to have some information on the price elasticity of supply. According to economic theory, information on the price elasticity of supply for a given commodity, its complements and competitors can be used as a tool for influencing supply response. As such, there is a need to analyze the supply response of sugarcane farmers to changes in its own price, the prices of a competing enterprise (maize) in terms of available land for cultivation and other policy measures especially in the face of impending withdrawal of COMESA sugar safeguards as they have been found to fuel producer and processor inefficiencies leading to low competitiveness of Kenyan sugar in the regional markets (Odhiambo, 2008).

## 1.2 Problem statement

Many crop-level supply response studies in Kenya have mainly focused on maize (e.g., Mose, 2007; Olwande, 2008) and other annual crops like potatoes (Muigai, 2013). However there is no recent study undertaken specifically on sugarcane supply response. Recent studies on sugarcane were on efficiency and total factor productivity (Mulwa *et al.,* 2005; Gicheru *et al.,* 2007). Since Mbogoh (1980), no recent study has focused on the sugarcane supply response to prices and other policy changes in Kenya.

Mbogoh (1980) assessed the factors that affect the supply of sugarcane and estimated the supply response of sugarcane farmers to price and non-price factors. In 2001, however, there was a major policy shift in Mumias Sugar Company when it was privatized and majority shareholding transferred from the government to private companies and the general public. Despite the fact that privatization has been encouraged as the way to steer the sub-sector to growth (Odhiambo, 2008), there is no study that has addressed the impact of privatization on the supply response of sugarcane farmers in Kenya. While the objectives of this study were similar to Mbogoh (1980), the current study goes further to estimate responsiveness of sugarcane farmers in the presence of privatization. This study provides an empirical understanding of the factors that influence the supply response of farmers that supply sugarcane to Mumias Sugar Company.

## 1.3 Purpose and objectives

The purpose of this study was to examine the responsiveness of outgrowers that supply Mumias Sugar Company with sugarcane to price, institutional factors and other economic incentives. The specific objectives were to:

(i) Characterize sugarcane production among outgrowers supplying cane to Mumias Sugar Company with emphasis on institutional issues.

(ii) Assess the responsiveness of outgrowers supplying cane to Mumias Sugar Company to price (such as own price and price of a competing enterprise, maize) and non-price factors.

## 1.4 Hypotheses

The hypotheses tested in this study were that:

1) Sugarcane price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run.
2) Maize price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run.
3) Non-price factors such as growth in technological advancement and yield taken singly have no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run.
4) There is no difference between the supply responsiveness of sugarcane farmers contracted by Mumias Sugar Company before and after privatization.

## 1.5 Justification of the study

Value addition is identified in the Kenya Vision 2030 as one of the main ways to steer growth in the agriculture sector. However, it can be greatly hindered if the existing industries operate below full capacity due to lack of sufficient raw materials. This study provides some vital information on the supply response of sugarcane farmers to own-price and prices of a competing enterprise (maize), which could contribute to the design of policies aimed at improving competitiveness of the sugar industry in Kenya. Understanding the factors that influence the supply response of farmers is important for proper and efficient policy formulation. Estimating the responsiveness of cane producers to price and non-price factors can help to identify the factors influencing the behavior of sugarcane farmers, which could improve the performance of Kenya’s sugar industry.

This study was carried out on Mumias Sugar Company, because the company contributes over 60 percent of the country’s domestic sugar production and hence is very vital to the sub-sector. The study specifically focused on outgrowers who provide more than 80 percent of the raw materials to the milling company. Understanding what influences the behavior of the farmers supplying Mumias Sugar Company can help in the formulation of policies geared towards improving the sub-sector as a whole. Additionally, Mumias Sugar Company was privatized in 2001 and hence it is possible to draw lessons from the behavior of the sugarcane farmers in the presence of privatization. This is particularly important because under the economic pillar, the Kenya Vision 2030 identifies the transformation of key institutions in agriculture as one of the ways to achieve growth in agriculture and hence overall economic growth. Besides, the privatization of the government owned sugar factories is one of the conditions related to the granting of import safeguards within COMESA (USDA, 2014).

The study is of interest mainly to policy makers, farmers, and other stakeholders including COMESA. For the policy makers, the study generated evidence based findings which can be used for efficient policy formulation and strategic planning. The findings can also be used to generate farmer advice for informed decision making. Its findings will help inform decision making on appropriate policies for long-term planning of sugar production in order to attain the goal of self-sufficiency. It is also a valuable addition to the existing stock of knowledge on supply response analysis.

## 1.6 Organization of the thesis

This thesis is organized as follows: Chapter 1 presents background information on the Kenyan sugar industry, statement of the problem, purpose and objectives of the study and finally justification of the study. Chapter 2 undertakes a review of the relevant literature; both theoretical modeling approaches and empirical studies are presented. Chapter 3 details the methodology, including the analytical framework and empirical models employed. The chapter also gives a detailed description of the data used in the analysis and the data analysis process. The results are presented and discussed in Chapter 4, while the recommendations and conclusions are presented in Chapter 5.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Theoretical review

### 2.1.1 Agricultural supply response concept

Agricultural supply response is a term used to describe the degree to which production changes due to changes in some important variables such as output price, scale of production and prices of substitutes. The concept attempts to explain behavioral response of producers to changes in economic incentives (Nkang *et al.*, 2007). For instance, the degree of responsiveness of farmers to changes in prices for a particular commodity is measured by the own-price elasticity of supply for the commodity. Own-price elasticity of supply is defined as the proportionate change in quantity supplied resulting from a percentage change in its price (Kiiru, 2006). According to economic theory, own-price elasticity of supply for a normal good should be positive (Nicholson and Snyder, 2008).

Reliable estimates of supply are important for predicting farmer responsiveness to input-output prices and thereby for formulating successful agricultural incentive programmes consistent with national goals (Ullah *et al.*, 2012). If the main objective of agricultural price policy is to stimulate production under existing technology, the policy is only relevant when farmers react rationally to price changes (Kiiru, 1995). A price policy is deemed to have a relatively low effect on output if the own-price elasticity of supply is low (Haughton, 1985). The purpose of supply response analysis is to examine how output responds to changes in factors that influence supply and are amenable to manipulation through policy. The importance of supply response studies is anchored on the fact that they give insight on how different policies affect production decision making.

### 2.1.2 Modeling agricultural supply response

The estimation of farmer supply response to price and other incentives was first undertaken by Nerlove in his seminal work in 1958 (Askari and Cummings, 1977). Since then, numerous other studies have been undertaken to examine the responsiveness of agricultural producers to price and non-price factors that affect production decisions and hence supply. While supply response studies can be done for broad agricultural aggregates of different commodities (Mythili, 2008; Muchapondwa, 2009; Obayelu and Salau, 2010), those for single commodities are more insightful for the formulation of appropriate sub-sector policies (Mose, 2007; Olwande *et al.*, 2009; Ozkan *et al.*, 2011).

Sadoulet and de Janvry (1995) indicate that there are three variants of agricultural supply response namely; production (output) response, acreage response and yield response. The choice between estimating output and acreage response mainly depends on the kind of data available. Output is mainly used when there is insufficient data on acreage or where acreage is constant over time (Mose, 2007). Nerlove (1956), however, argues that the area under crop is the only variable directly under the control of the farmer and hence gives preference to the use of acreage response as proxy for supply response. While the argument may be partly right, the application of yield enhancing inputs like fertilizers are also to a large extent directly under the control of the farmer. Proponents of the acreage response approach argue that output and yield are more susceptible to the effects of uncertain random factors such as rainfall and temperatures (Mythili, 2008) than area. Therefore, output or yield may not actually represent the true response of farmers to price and non-price factors in the production environment (Nerlove, 1956). To understand the behavioral pattern of farmers, therefore, area is a more appropriate variable.

In the estimation of area response, some studies use the ratio of crop acreage to total area cultivated as the dependent variable. However, in the event that there are simultaneous changes in crop area and total area cultivated, the variations may be concealed because the ratio remains the same or changes very marginally (Bingxin *et al.,* 2010). To counter this problem some analysts prefer the absolute area as the dependent variable.

There are two broad approaches in literature to supply response analysis that can be employed in the estimation of area, output or yield response (Mythili, 2006). The first is the derivation of supply functions through the profit maximization approach while the other is the use of dynamic models. The former involves the joint estimation of output supply and input demand functions derived from the Hotelling's lemma and requires detailed information on all input prices (Mythili, 2008). As such, it is mostly used with cross-sectional data because it is generally difficult, particularly in developing countries, to get time-series or balanced panel data with all the required variables (Olwande *et al.,* 2009). Ozkan *et al*. (2011) refers to this approach as "direct structural" analysis. The approach makes inference from the equilibrium status of input demand and output supply functions and relates them to the production function (Ozkan *et al.,* 2011).

The nature of agricultural production however, is such that the response to supply shifters is not instantaneous (Muchapondwa, 2009). First, agricultural supply response is characterized by biological lags between input application and output production. Second, there exist technological and institutional factors that prevent intended production decisions from being fully realized in any one period. Third, the assumption of perfect information cannot apply to agricultural production because the environment is characterized by information asymmetries especially on prices (Muchapondwa, 2009). These reasons necessitate the use of dynamic models where a time variable is introduced to capture these attributes.

Some of the commonly used dynamic supply response models are distributed lag and autoregressive models. Distributed lag models include lagged explanatory variables, meaning that the effect of a unit change in the explanatory variable is distributed over a number of time periods (Gujarati, 1999). Autoregressive models, on the other hand, include lagged values of the dependent variable. Models that include a lagged dependent variable as well as lagged explanatory variables are referred to as Autoregressive Distributed Lag Models (ADL) (Gujarati, 1999).

Following Kapuya (2010), a distributed lag supply model is based on the premise that the quantity supplied in period *t* is a function of the price received in the previous period *t-1*, that is:

$$\boldsymbol{Q_t = \alpha + \beta p_{t-1} + \mu_t} \qquad \textbf{(2.1)}$$

where: $Q_t$ is the quantity supplied in period $t$

$\alpha$ is the intercept

$p_{t-1}$ is the price paid in period $t-1$

$\beta$ is the short run elasticity that measures the degree of responsiveness of $Q$ to a unit change in $p_{t-1}$

$\mu_t$ is the error term

Because the effect of the changes in one variable may be felt through several periods, this leads to a distributive lag equation which is specified as follows (Kapuya, 2010):

$$\boldsymbol{Q_t = \alpha + \beta_0 p_t + \beta_1 p_{t-1} + \cdots + \beta_k p_{t-k} + \mu_t} \qquad \textbf{(2.2)}$$

where: $p_{t-k}$ is the price received in time period $k$, $\forall\ k = 2, ..., n$

$n$ is the maximun number of lags

$\beta$ are parameters to be estimated and the other variables are as defined in the previous equation.

The model in Equation (2.2) usually experiences two major problems. The first one is multicollinearity because successive values of economic variables tend to be serially correlated. Using the ordinary least squares (OLS) technique would lead to biased and inefficient parameter estimates. If the data are tested and found to be stationary or corrected for stationarity, OLS can still be used (Kapuya, 2010). The second problem is that it is difficult to decide the number of lagged values of explanatory variables to introduce (Gujarati, 1999). If too many lagged values are included, a problem of degrees of freedom is introduced. This makes the results from the model unreliable because with few degrees of freedom the analyst cannot confidently deduce that the sample adequately represents the population under study. To reduce the number of lagged variables and deal with the issue of multicollinearity in distributed lag models, the Koyck transformation process[3] is applied to come up with a simplified version of the model, as specified in Equation (2.3):

$$\boldsymbol{Q_t = \alpha + \beta_0 p_t + \beta_1 Q_{t-1} + \mu_t} \qquad \textbf{(2.3)}$$

where $Q_{t-1}$ is the quantity supplied in period *t-1* while the other variables are as previously defined in equations 2.1 and 2.2. Because the model contains a lagged value of the dependent variable ($Q_{t-1}$), it is generally referred to as an autoregressive model. Such models are commonly used in dynamic single-equation modeling, for instance, in adaptive expectations and partial/stock adjustment models (Hassler and Walters, 2006).

The Nerlovian model is based on partial adjustment and expectations for prices and quantities. The model specifies the production function with an equation formed out of the price expectations and non-price supply shifters (Ozkan *et al.*, 2011). It is based on the assumption that the quantity supplied in year *t* is a linear function of the expected price for the year.

$$\boldsymbol{Q_t^* = \alpha + \beta p_t^* + \beta Z_t} \qquad \textbf{(2.4)}$$

where $Q_t^*$ is the desired level of output for year *t*

$p_t^*$ is the expected price for year *t*

$Z_t$ is a vector of non-price exogenous variables.

[3] Koyck transformation is a process used to transform an infinite geometric lag model into a finite model with lagged dependent variable. See Gujarati (1999), pp 441-442 for details.

The expectations theory assumes that farmers make production decisions for the current period in previous period(s) and cannot revise their decisions in the current period (Ozkan *et al*., 2011). As such, the difference in output levels between two time periods is theoretically a function of the difference between the desired and actual output levels as expressed in Equation (2.5):

$$\boldsymbol{Q_t - Q_{t-1} = \delta(Q_t^* - Q_{t-1})} \qquad \textbf{(2.5)}$$

where: $Q_t$ is quantity supplied in period *t*

$Q_{t-1}$ is the output level supplied in period *t-1*

$Q_t^*$ is the desired output level and

δ is the proportion of the difference in output level between the time periods explained by the difference between $Q_t^*$ and $Q_{t-1}$

The model also assumes that the expected price in period *t* equals the expected price in period *t-1* plus a fraction of the difference between actual and expected price in period *t-1* (Muchapondwa, 2009). This relationship can be expressed as:

$$\boldsymbol{p_t^* = p_{t-1}^* + \delta(p_{t-1} - p_{t-1}^*)} \qquad \text{where } 0 \leq \delta \leq 1. \qquad \textbf{(2.6)}$$

The difference between actual and expected price in period *t-1* represents the error made in the previous season in attempting to predict prices for the next period (Kapuya, 2010). Rearranging the Equation (2.6) yields:

$$\boldsymbol{p_t^* = \delta p_{t-1} + (1 - \delta)p_{t-1}^*} \qquad \textbf{(2.7)}$$

Inserting Equations (2.7) and (2.5) into Equation (2.4) and rearranging yields a supply response function based on adaptive expectations represented in equation 2.8 (Kapuya, 2010):

$$\boldsymbol{Q_t = \alpha + b\delta P_{t-1} + \ 1 - \delta \ Q_{t-1} + b\delta Z_t} \qquad \textbf{(2.8)}$$

Equation (2.8) now contains observable variables and hence can be easily estimated. If the data are in logs, the coefficient on $P_{t-1}$ which is *b* δ is the short-run price elasticity. The long-run elasticity, *b,* can be obtained by dividing *b* δ by δ while δ is obtained by subtracting the coefficient on the lagged dependent variable $Q_{t-1}$ from 1, i.e., 1-(1- δ). Note that 1- δ is a parameter obtained directly from the regression. Similarly, the coefficient of each explanatory variable directly gives short run elasticities, while the long run elasticities are obtained by dividing the short run elasticities by [1- coefficient of the lagged dependent variable] (Mythili, 2008).

The adaptive expectations and partial adjustment models are the two variants of the Nerlove supply model (Mose, 2007), which is a widely used dynamic approach for supply response analysis (Ozkan *et al.,* 2011). These models however, cast producers as being very mechanical in making production decisions and incapable of learning from previous errors hence reducing supply response analysis to an *ad hoc* assumption that in each time period, a fraction of the difference between current and the long-run is eliminated (Mose, 2007). Another criticism is that they are based on price weights that decline geometrically. Additionally, these price weights are subjective, and not based on the explicit outcome of an optimization process (Olwande *et al.,* 2009). Both the adaptive expectations and partial adjustment models employ the OLS technique to estimate a dynamic specification of supply response (Mose, 2007). This implies that the estimates are based on the underlying assumption that the data processes are stationary because the data are regressed in their level[4] form (Mose, 2007). Conducting time series analysis using data that have unit root leads to spurious and invalid results (Greene, 2003).

Most predictions in classical economic theory are based on the assumption that observed data come from a stationary process where the means and variance are constant over time (Hendry and Juselius, 2000). When this is not the case, the data are said to have a unit root (i.e., are non-stationary) and follow a random walk (Maddala, 2002; Greene, 2003). However, most

[4] Level data refers to a series in its original form before differencing (Mose, 2007).

economic variables, including agricutural time series, are non-stationary (Mose, 2007). As such conventional regression analysis is not encouraged for the estimation of area or output response (Clark and Clein, 1996) because it ignores the non-stationary properties of time series data. Additionally, regressing time series that are trended overtime may produce high goodness-of-fit statistics suggesting that the data are highly correlated yet the relationship is mostly spurious (Ozkan *et al.*, 2011).

When the data series are found to be non-stationary analysts often resort to the use of Vector Autoregression (VAR) and Vector Error Corection Models (VECM) which are in the class of ADL models (Maddala, 2002; Gujarati, 1999). It is however important to note that including many lags of the independent variables could lead to multicollinearity (Gujarati, 1999). The VAR model is a general framework used to analyze the dynamic inter-relationship among variables that are either stationary or non-stationary (Muchapondwa, 2009). A general VAR is appropriate to analyze data series that are not co-integrated[5]. When the variables are co-integrated, however, some adjustments to the VAR are necessary and hence the need for co-integration analysis. The ADL approach to co-integration and the VECM are the most commonly used methods of co-integration analysis (Muchapondwa, 2009).

The ADL approach to cointegration is used when the data series are co-integrated but have differing levels of integration (Muchapondwa, 2009). For example, when some variables are stationary, i.e., I(0) while others are integrated of order 1, i.e., I(1). A VECM, on the other hand, is a variant of the VAR for variables that are non-stationary and co-integrated of the same order. This type of model directly estimates the rate at which changes in the dependent variable return to equilibrium after a change in an independent variable (Gujarati, 2004). It is justified in that it implies that the behaviour of the dependent variable is tied to the changes in

[5]Co-integration means that the residuals from a relationship between variables are stationary and normally distributed (Pfumayaramba, 2011).

the explanatory variables in the long-run and that changes in the dependent variable respond to deviations from that long-run equilibrium.

To illustrate the concept of non-stationary series and co-integration, consider the relationship between quantity supplied and price specified as follows (Ozkan *et al.*, 2011):

$$\boldsymbol{Q_t = \alpha + \beta p_t + \mu_t} \qquad \textbf{(2.9)}$$

When $Q_t$ and $p_t$ are integrated, it is possible to have a linear equation that explains the relationship between the two co-integrated variables. Taking the first differences of the levels data to correct for unit root yields the following:

$$\Delta \boldsymbol{Q_t = \alpha_0 + \alpha_1} \Delta \boldsymbol{p_t + \gamma \mu_{t-1}} \qquad \textbf{(2.10)}$$

When Equation (2.9) is rearranged to make the error term the subject and substituting into (2.10) gives:

$$\Delta \boldsymbol{Q_t = \alpha_0 + \alpha_1} \Delta \boldsymbol{p_t + \gamma}(\boldsymbol{Q_{t-1} - \alpha - \beta p_{t-1}}) \qquad \textbf{(2.11)}$$

Because $\Delta Q_t = Q_t - Q_{t-1}$ and $\Delta p_t = p_t - p_{t-1}$, rearranging yields:

$$\boldsymbol{Q_t = \ \ \alpha_0 + \gamma\alpha \ \ + \alpha_1 p_t - \alpha_1 p_{t-1} + \ \ 1 - \gamma \ \ Q_{t-1} + \gamma\beta p_{t-1}} \qquad \textbf{(2.12)}$$

Equation (2.12) can be re-written as:

$$\boldsymbol{Q_t = \ \ \alpha_0 + \gamma\alpha \ \ + \alpha_1 p_t + \ \ 1 - \gamma \ \ Q_{t-1} + \ \ \gamma\beta - \alpha_1 \ \ p_{t-1}} \qquad \textbf{(2.13)}$$

The transformations lead to the Granger representation[6] in Equation (2.13), which models the short-run and long-run adjustments and overcomes the impediments of the Nerlovian expectations and adjustment models (Ozkan *et al.*, 2011). The VECM derives from the basic framework of the Granger representation.

A simple formulation of an ECM is as follows:

$$\Delta \boldsymbol{A_t = \beta_0} \Delta \boldsymbol{X_t + \ \gamma \ A_{t-1} - \beta_1 X_{t-1} \ + \ \omega_t} \qquad \textbf{(2.14)}$$

[6] The Granger representation theorem states that a co-integrated vector autoregressive process can be decomposed into four components: a random walk, a stationary process, a deterministic part, and a term that depends on the initial values (Hansen, 2005).

where $\Delta$ is the first difference operator, i.e., $\Delta A_t = A_t - A_{t-1}$. $B_0$ captures the short-run relationship between $A$ and $X$. It indicates how $A$ and $\Delta A$ immediately change if $X$ goes up by a unit in one period. The term $A_{t-1} - \beta_1 X_{t-1}$is a formulation to show that it is assumed that $A_t$ and $X_t$ have a long-run equilibrium relationship. $\gamma$ is the error correction term which shows the rate at which the model returns to equilibrium. In other words, $\gamma$ is the proportion of the disequilibrium which is corrected within one time period. The error correction term should be negative. If $\gamma$ =0, it means that the model does not revert back to equilibrium after a shock while $\gamma$=-1 means that equilibrium is achieved after one time period (Gujatari, 2009).

The ECM methodology allows for the formulation of a model out of non-stationary variables and enables statistical inference without imposing restrictions on the short-run behaviour of the variables (Ozkan *et al.*, 2011). Its major advantage over the ADL model is that it incorporates both short-run and long-run effects without imposing restrictions on the behaviour of the variables. The error correction mechanism is based on the coefficient of the lagged dependent variable in an ADL model augmented with explanatory variables and stems from a stable long-run relationship of the variables (Mohammad, 2007).

The most widely used single equation approach to co-integration is the Engle-Granger two-step procedure. This approach ignores short-run dynamics when estimating the co-integrating vector and thus biases the estimate of the long-run relationship in finite samples (Muchapondwa, 2009). To overcome this problem, a test based on the coefficient of the lagged dependent variable in an ADL framework proposed in Banerjee *et al*. (1998) is used. However, the procedure assumes that only one co-integrating vector exists, which leads to inefficiency in estimation in the event that more than one co-integrating vector actually exists. The Johansen estimation procedure counters this limitation by allowing for the establishment of the number of co-integrating relationships (Mohammad, 2007).

## 2.2 Specification of functional forms in supply response estimation

There are various specifications of functional forms in literature that can be adopted for supply functions. Among these are the linear, semi-log and double log functional forms (Gujarati, 1999). When the functional form is linearized by taking logarithms, elasticities are obtained directly from the model as parameter estimates for the respective variables (Ozkan *et al*., 2011). Additionally, this does away with the need to refer to the units of measurement of the variables in the regression (Greene, 2010). When there are more than one explanatory variables in the model, a multivariate double log functional form is applied. In the double log functional form, each partial derivative measures the elasticity of respective explanatory variables on the dependent variable while holding the effects of other variables constant (Gujarati, 1999).

## 2.3 Empirical studies on supply response

Several supply response studies have been undertaken for different agricultural commodities in and outside Kenya. Mose (2007) used co-integration and VECM approaches to estimate the supply response of maize in Trans-Nzoia District for the period 1980 to 2002. The study found that farmers respond strongly to price incentives as evidenced by the significant but negative short- and long- run elasticity of supply to fertilizer price and positive own-price elasticity of supply in the short- and long- run. The study further tested for differences in the supply response in the pre- and post-liberalization era. No differences in both short-run and long-run supply responses between the two periods were found. The study concluded that the supply response could not be wholly attributed to market liberalization. The current study benefits from Mose (2007) in terms of methodological approach, especially with regard to the use of VECM and the inclusion of a structural break variable which helped to gauge whether there were any differences in supply response during different policy regimes.

Olwande *et al*. (2009) assessed the responsiveness of maize production to price and non-price factors in the high potential maize growing areas of Kenya. Employing a normalized restricted translog profit function, the authors estimated maize supply and variable input demand elasticities using cross-sectional data from a sample of farmers in Kenya's high potential maize zone. The study found that the area of land allocated to maize and fertilizer use were important factors affecting maize supply in Kenya. The study concluded that maize price support would be an inadequate policy for expanding supply in the country. This is because a price support policy would result in higher consumer prices, and hence only benefit a small section of the producers who are commercial, but hurt the welfare of the majority who are either smallholder producers or urban consumers. Although Olwande *et al.* (2009)'s study was done for a different crop, using cross-sectional data and used a different type of model, it is informative to this study in terms of providing some of the important variables to include in the estimation of a supply response model.

Zulfiqar *et al*. (2011) assessed the determinants of sugar supply and demand in Pakistan using a simultaneous equation recursive model. The domestic price of sugar, fertilizer use and water availability were found to positively or negatively influence domestic supply. The study recommended the formulation of policies geared towards making fertilizers accessible to farmers; government investment in the provision of irrigation water, and research and extension services in order to optimize resource use in sugarcane production in Pakistan. Zulfiqar *et al*. (2011)'s study informed this study by providing important background information on the determinants of sugarcane supply response.

Using the ADL approach to co-integration, Muchapondwa (2009) examined the supply response of Zimbabwean agriculture under different pricing regimes. The study used time series data for the period 1970 to 1999 to estimate the aggregate agricultural supply response to price and non-price factors. The long-run supply was found to be price inelastic (0.18), implying that the pricing policy was not effective in fostering growth in aggregate agricultural supply. The study recommended the provision of non-price incentives to revive the agricultural sector in the country. The ADL approach to cointegration used in Muchapondwa (2009) is used when the variables have differing levels of integration.

Obayelu and Salau (2010) estimated the responsiveness of Nigerian agricultural supply to prices and exchange rates. Employing co-integration and VECM on time series data spanning from 1970-2007, the authors found that prices and exchange rates were significant in determining supply both in short- and long- run. The study concluded that changes in crop prices and exchange rates were affecting supply. The study provides a guideline on some of the tests needed on time series data to improve the robustness of the results. As such, the current study borrows a lot from Obayelu and Salau (2010) in terms of the methodological approach.

Kapuya (2010) studied the impact of ‘fast track’ land reform policy on Zimbabwe’s maize production using a partial equilibrium model that depicted what would happen if the policy shift had not occurred. The study found that the size of the commercial area harvested was negatively affected by the policy. It concluded that the total area harvested would have been higher under the pre-2000 policy conditions. The land reform policy shift led to a decline in area as a result of the stalling of farm operations due to political unrest, economic instability and input shortages. The study used the Nerlovian approach to supply response analysis, which informed the current study on the considerations for various methodological approaches. It is an example of the study of a policy shift that failed to attain the intended results in a developing country context.

Mesike *et al.* (2010) evaluated the supply response of rubber farmers in Nigeria for the period 1970 to 2008. The study used co-integration and VECM techniques and found that producer prices and a policy shift in the form of Structural Adjustment Progammes significantly affected the supply of rubber. The elasticities were, however, low with magnitudes of 0.37 and 0.20 in the short-run and long-run periods respectively. The authors concluded that rubber farmers were not very responsive to prices owing to the emergence of other supply determinants. The study recommended policy efforts to promote sustainable marketing outlets which offer better producer prices. The current study benefited from Mesike's study with regard to methodological approaches since it captures the effect of privatization on sugarcane supply.

## 2.4 Synthesis

Among the supply response studies reviewed, some used cross-sectional farm level data to estimate profit functions leading to output supply and factor demand functions. While the profit function approach is considered econometrically sound, it often underestimates elasticities due to its inability to take into account the time required by producers to adjust their production decisions (Olwande *et al.*, 2009). The elasticity estimates from this approach may therefore not reflect the true responsiveness of agricultural producers.

Other studies employed dynamic models using the Nerlovian, ADL approach to cointegration and VEC co-integration approaches. The Nerlovian approach is more commonly used for studies covering broad agricultural aggregates using country level data, however, it is also used for single commodities. The Nerlovian approach suffers from the tendency to underestimate elasticity due to its inability to differentiate between short- and long-run supply response parameters (Piya, 2002; Muchapondwa, 2009) and capture the full dynamics of agricultural supply due to the imposition of many restrictions. Cointegration analysis derives distinct parameters for short- and long-run estimates. The ADL Approach to cointegration takes into account the possibility of reverse causality (Muchapondwa, 2009) and is appropriate for series that are co-integrated of different order. The current study used the VECM as the most appropriate method when the data series are non-stationary and co-integrated of the same order.

# CHAPTER 3: METHODS AND DATA

## 3.1 Theoretical framework

Supply response analysis is anchored on the theory of the firm which postulates that under perfect market conditions producers are profit maximizers (Varian, 1992). Thus, a profit maximizing firm will choose its output level by equating its marginal cost to the marginal revenue (Nicholson and Snyder, 2008). When the producer (e.g., farmer) is a price taker, the profit maximizing level of production equates the marginal cost to the output price. In the short-run therefore, the firm's supply function is the upward sloping portion of the short-run marginal cost curve where the price is equal to or greater than the minimum average total cost (Varian, 1992).

According to the theory of the firm, the market supply of a commodity is determined by its own price, the price of substitutes and complements, as well as those of inputs and technology (Piya, 2002; Muchapondwa, 2009). Other determinants of supply are firm's expectations about future prices and the number of firms supplying the market. The responsiveness of firms to changes in the producer environment is measured by the elasticity of supply. The short-run price elasticity of supply measures the responsiveness of suppliers in an industry to changes in prices. This measure shows how proportionate changes in market price are met by changes in total output (Nicholson and Snyder, 2008).

The assumptions of a perfectly competitive market however, do not always apply in many industries. The Kenyan sugarcane sub-sector can be classified as largely oligopsonistic because many cane farmers supply few sugar milling companies. No single farmer can influence the producer price of sugarcane and hence the cane farmers are price takers.

Apart from profit maximization, the other consideration in the analysis of the firm is the biological lag and other technological and institutional factors that impede farmers from adjusting output instantly after a change in prices and other factors (Piya, 2002). An agricultural firm may therefore not behave like a typical firm. Based on the two problems: profit maximization and the biological lag problem, there are two theoretical frameworks that are used to model supply response in agricultural firms. One is the Nerlovian expectation model and the other is the supply function derived from Hotelling's lemma based on the profit maximizing framework. The supply function requires detailed information on all input prices. In a developing country set up however, such detailed data may not be available.

The biological lag problem is handled in a dynamic model specification. Following Nerlove (1958), farmers form price expectations based on the most recent season's prices. The model however assumes that the data generation process is stationary and therefore fails to take into account the unit root problem (Piya, 2002). This often leads to spurious regression problem when the data are not stationary (Mose, 2007). The ECM resolves the problem of non-stationarity and analyzes the relationship under a dynamic approach.

The first step in time series analysis is to determine whether the variables are stationary or not. A variable is said to be integrated of order $n$ [i.e., I($n$)] if unit roots or stochastic trends can be removed by differencing the variable $n$ times and a stochastic trend still remains after differencing $n-1$ times (Lütkepohl, 2005; Greene, 2003). If a variable does not have a stochastic trend or a unit root, it is stationary and referred to as I(0). Therefore, to make a series that has unit root stationary, all the analyst needs to do is to difference the data. Alternatively, this can be achieved through the inclusion of a trend or both. As such, the series becomes difference stationary or trend stationary, respectively. This process of transforming data into stationary series (differencing and detrending) however leads to loss of some valuable long-run information, a problem that can be solved through an Error Correction Model (ECM) framework.

The linear transformation of differencing makes data that has unit root to be stationary. If the data generation process is a simple random walk whose error term is independent and normal (IN) with mean zero and a constant variance, i.e., $X_t = X_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim \mathrm{IN}[0, \sigma_\varepsilon^2]$, then subtracting $X_{t-1}$ from both sides of the equation gives $\Delta\varepsilon_t \sim \mathrm{IN}[0, \sigma_\varepsilon^2]$ which is stationary (Dickey and Fuller, 1979). It follows that differencing data that is I(2) twice makes it I(0). This deduction can be made for whatever order of integration (Mose, 2007). A stationary series is said to be integrated of order zero, which is expressed as I(0). The extent of integration in a series is derived from the number of times the series needs to be differenced to become stationary (Greene, 2003). If the series becomes stationary after being differenced once then it is said to be integrated of order one, i.e., I(1).

Consider two time series variables *x* and *y* at time *t.* Expressing the two in a dynamic inter-relationship in which each variable is a function of its own lag and the lag of the other variable yields the following system of equations:

$$y_t = \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} x_{t-1} + \mu_t^y \qquad \textbf{(3.2)}$$

$$x_t = \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} x_{t-1} + \mu_t^x \qquad \textbf{(3.3)}$$

This system of equations is called a Vector Autoregression (VAR) (Hill et al., 2011). A VAR with maximum lag order one (like in the system above) is referred to as a VAR(1). If the series are not stationary, then the VAR needs to be modified to allow a consistent estimation of the relationship. This is done by taking the first differences of the series and testing the resultant data for stationarity (Mose, 2007).

$$\Delta y_t = \beta_{10} + \beta_{11} \Delta y_{t-1} + \beta_{12} \Delta x_{t-1} + \mu_t^{\Delta y} \qquad \textbf{(3.4)}$$

$$\Delta x_t = \beta_{20} + \beta_{21} \Delta y_{t-1} + \beta_{22} \Delta x_{t-1} + \mu_t^{\Delta x} \qquad \textbf{(3.5)}$$

If *y* and *x* are not stationary in their levels but stationary in the first differences, expressed as I(1), then the process involves taking their first differences and using OLS to estimate the model. If however the variables are I(1) and co-integrated, then the system should be modified to allow for the co-integrating relationship. The resultant model is the Vector Error Correction Model (VECM) which is a special case of the VAR for variables that are stationary in their differences and co-integrated (Obayelu, 2010).

According to Pfumayaramba (2011), any linear combination of I(1) variables is spurious. However, if there is a long-run relationship, the errors tend to revert back to zero, i.e., I(0). The spurious regression problem implies that regressing independent non-stationary variables can result in evidence of a relationship when none exists. To overcome this limitation, there needs to be evidence of co-integration among the varibles. If two I(1) series, *X* and *Y*, have a relationship such that the residuals from the regression, $Y_t=b_0+b_1X_t + \mu_t$ are stationary, then the variables are said to be co-integrated (Pfumayaramba, 2011). This means that if a linear combination of I(1) variables which is I(0) exists then, the variables are co-integrated. As indicated earlier, co-integration implies that the residuals from a relationship between variables are stationary and normally distributed (Pfumayaramba, 2011). It follows that if variables are non-stationary then it is only possible to infer a long-run relationship if they are co-integrated.

The estimation of a plausible VECM requires that the data series used in the analysis be co-integrated (Mose, 2007). This is because co-integration among a set of variables implies that there exist fundamental economic forces, which make the variables to move stochastically together over time. The ECM then corrects for any disequilibrium between variables that are co-integrated because the sequence of the discrepancy between the observed and the equilibrium state tends to revert back to its mean, which is zero.

The Engel-Granger test for co-integration is widely used to test the residuals from the model for non-stationarity. If two series $y_t$ and $x_t$ are co-integrated, there exists a linear relationship between them that is stationary. Consider the relationship between two variables *y* and *x* such that: $y_t - \beta x_t = \mu_t$. If $\mu$ was known, it would be directly tested for stationarity using the Dickey-Fuller test. But because it is not known and $\beta$ are also unknown, the first step is to run an OLS regression on the series and determine $\beta$, then estimate $\mu$ in the second step and test them for stationarity (Sjo, 2008).

For example, in the relationship, $Y_t = bX_t + \mu_t$, the co-integration test equation will be as follows:

$$\Delta\varepsilon_t = \rho * \varepsilon_{t-1} + \sum_{i=1}^{k} \rho_i * \Delta U_{t-i} + \omega_t \qquad \textbf{(3.6)}$$

$H_0 : \rho^* = 0$, i.e., no co-integration exists (non-stationary residuals)

$H_1 : \rho^* < 0$, i.e., cointegration exists (stationary residuals).

Where $\rho^*$ is the test coefficient,

$\Delta$ is the difference operator,

$\varepsilon_t$ and $\varepsilon_{t-1}$ are the values of the error term in period *t* and *t-1* respectively.

$\Delta U_{t-i}$ is change in residual values in time period *i*, $\forall$ *i=2...... k*

If the test result is such that the $H_0$ is rejected, co-integration is confirmed and this implies that the residuals are stationary and hence an ECM can be used for the analysis (Pfumayaramba, 2011).

The limitation of the Engel-Granger test for co-integration is that it can only be used for two time series variables. Additionally, the Engel-Granger approach estimates a co-integrating parameter from a static regression equation which could be susceptible to small sample bias (Olubode-Awosola *et al*., 2006). The other limitation of the Engel-Granger test for co-integration is that it assumes a single co-integration vector. The Johansen test, on the other hand, allows for the empirical determination of the number of co-integrating vectors in more than two time series variables (Thiele, 2002; Kuwomu *et al*., 2011).

The Johansen test uses the reduced rank regression procedure (Nkang *et al*., 2007). It determines the number of co-integration vectors using two tests: the maximum eigenvalue test and the Trace test (Asari *et al.*, 2011; Buigut, 2011). The maximum eigenvalue statistic tests the null hypothesis of *r* co-integrating relations against the alternative hypothesis of *r+1* co-integrating relations for *r* = 0, 1, 2, ........, *n*-1. The statistic is computed as:

$$\boldsymbol{LR_{max}}\,(\boldsymbol{r}\,/\,\boldsymbol{n}\,+\mathbf{1}) = -\boldsymbol{T} * \mathbf{log}(\mathbf{1}-\boldsymbol{\lambda}) \qquad \textbf{(3.7)}$$

where λ is the maximum eigenvalue and *T* is the sample size.

The trace statistic, on the other hand, tests the null hypothesis of *r* co-integrating relations against the alternative of *n* co-integrating relations where *n* is the number of variables in the system for *r* = 0, 1, 2, ....., *n*-1. The static is computed as:

$$\boldsymbol{LR_{trace}}\ \boldsymbol{r}\ \boldsymbol{n} = -\boldsymbol{T} * \sum_{\boldsymbol{i}=\boldsymbol{r}+\mathbf{1}}^{\boldsymbol{n}} \mathbf{log}\ \mathbf{1}-\boldsymbol{\lambda}_{\mathbf{i}} \qquad \textbf{(3.8)}$$

For the results of a VAR or VECM to be reliable, it is necessary that the correct number of lagged dependent and other explanatory variables is included. According to Lütkepohl and Kratzig (2004), selecting a higher order lag length increases the mean square errors of the VAR/VECM while underestimating the lag length often generates autocorrelated errors.

In order to determine how many lags to use, several selection criteria can be used. The two most common are the Akaike Information Criterion (AIC) and the Schwarz' Bayesian Information Criterion (SIC/BIC/SBIC). These criteria choose lag length *j* to minimize: log(SSR(j)/n) + (j + 1)C(n)/n, where SSR(j) is the sum of squared residuals for the model with *j* lags and *n* is the number of observations; C(n) = 2 for AIC and C(n) = log(n) for BIC (Komeh, 2012; Lütkepohl and Kratzig, 2004; Greene, 2003). Lütkepohl and Kratzig (2004) argue that restrictions meant to improve the estimation precision in a VECM may be based on economic theory or other non-sample information and on statistical procedures.

## 3.2 Analytical framework

Based on the foregoing discussion, this study used the VECM, a variant of the VAR model for non-stationary variables that are co-integrated, to estimate the supply response of sugarcane production among outgrowers contracted to supply Mumias Sugar Company in Kenya for the period 1980 to 2013. The main reason for the choice of the model was that stationarity tests on the time series data collected for this study showed that all the series were non-stationary in their levels and co-integrated. This meant that the use of conventional regression analysis like OLS would yield unreliable results. The Johansen estimation procedure was adopted in this study. This study also adopted a multiple double log functional form because there were more than one explanatory variables under consideration.

### 3.2.1 Empirical model

The theoretical model given in equation 2.14 was specified as follows:

$$\Delta A_t = \beta_0 + \sum_{i=1}^{k-1} \beta_{i1}\Delta A_{t-i} + \sum_{i=1}^{k-1} \beta_{i2}\Delta p^{y}_{t-i} + \sum_{i=1}^{k-1} \beta_{i3}\Delta p^{s}_{t-i} + \sum_{i=1}^{k-1} \beta_{i4}\,\Delta y_{t-i} + \sum_{i=1}^{k-1} \beta_{i5}\,\Delta ST_t + \lambda\,[A_t - \alpha_0 - \alpha_1 p^{y}_{t-1} - \alpha_2 p^{s}_{t-1} - \alpha_3 y_{t-1} - \alpha_4 ST_t - \alpha_5 T_t] + \omega_t \quad (3.9)$$

where:

Δ represents the first difference operator for the respective variables.

***A**$_t$* is dependent variable representing the acreage of sugarcane planted in year *t*. This is the proxy for planned supply. Area is preferred to output because it is the variable directly under the control of the farmer and hence the best proxy for the farmers' response to changes in the production environment (Nerlove, 1956; Alemu *et al*., 2003).

$\boldsymbol{\beta}$ are the short-run supply parameters. They measure the effect of a percent change in the respective explanatory variables on the dependent variable in the short-run.

$\boldsymbol{k}$ is the maximum number of lags included in the model as determined by the data properties.

$\boldsymbol{\lambda}$ is the error correction mechanism that measures the speed of adjustment from short-run disequilibria to long-run steady state equilibrium. It measures the extent of correction of errors in the dependent variable and its expected sign is always negative (Asari, 2011).

α are the long-run coefficients for the various explanatory variables.

$p_t^y$ is the price of sugarcane in time period *t.* An increase in the price of a commodity was expected to lead to an increase in its supply (Nicholson and Snyder, 2008) and hence its expected sign is positive.

$p_t^s$ is the price of the competing enterprise, in this case maize, in time period *t.* The variable was meant to capture the opportunity cost of producing the competing enterprise (Alemu, 2003). Changes in the price of substitute enterprise are expected to cause a shift of the supply curve (Nicholson and Snyder, 2008).

$y_t$ is the yield of cane in time period *t*. The expected profitability of sugarcane is not only dependent on the expected price and costs of production, but also on the expected yield. The yield attained in the previous seasons is the best proxy for expected yield (Jaforullah, 1992). The inclusion of the yield variable $y_t$ was meant to capture the effects of changes in non-acreage inputs and other exogenous variables that affect productivity like rainfall and temperatures (Mythili, 2008). Time series information on input use, especially fertilizer application and seed cane was not available. In the expectations model, if the yield was good in a particular season farmers were expected to increase the land under the crop in subsequent seasons in anticipation for more profits resulting from better yields. While yield can be correlated with area, it has been argued that small holder farmers respond to price changes by reallocating land among competing enterprises as opposed to putting more land under cultivation since they are land constrained (Mythili, 2008). But in cases where increase in prices brings into cultivation lands that previously were left uncultivated, the average yield may reduce. In the case of Western Kenya the population has been high and most of the land already under cultivation hence no reason to believe that the average yields decreased due to cultivating larger chunks of land. There is, however, no adequate information to show if more marginal lands were brought into cultivation in response to higher output prices. These estimates should be interpreted with caution.

*T* is the time trend variable. The trend variable served as a proxy for variables that affect the dependent variable and were not directly observable but are highly correlated with time or those whose data may not be available over a long period of time. These include historical data on infrastructural development, applications of modern farming techniques and expenditure on agricultural research and extension services, among others (Alemu, 2003). Because it is not always possible to include these variables directly and individually in the acreage response function, this approach is used in most studies that use time series analysis (Kiiru, 1995; Mesike *et al.*, 2010). Failing to take into account the effect of such variables would imply model underspecification. The possibility of model overspecification is taken care of through post-estimation tests.

The term of the equation that is in square brackets represents the divergence from the long-run such that, when equilibrium holds in the long-run, the term $A_t - \alpha_0 - \alpha_1 p_{t-1}^{y} - \alpha_2 p_{t-1}^{s} -$

$\alpha_3 y_{t-1} - \alpha_4 T_t \ - \alpha_5 ST_t$ = 0. The coefficients for each of the variables in this term are the long-run supply response parameters for the respective variables. It is normalized by the error correction term obtained in the short-run, that is, $\lambda$.

$\omega_t$ is the stochastic error term. It is assumed to be independently and normally distributed with zero mean and constant variance.

$ST_t$ is a dummy variable representing the change of company ownership and management from government to private ownership, thus $ST_t$=1 if t>=2001, and 0 otherwise. The inclusion of this variable made it possible to capture the effect of exogenous variables which may result from transitions to new policy regimes (Alemu *et al.*, 2003).

The long-run supply equation was specified as follows:

$$A_t = \alpha_0 + \alpha_1 p^y_{t-1} + \alpha_2 p^s_{t-1} + \alpha_3 y_{t-1} + \alpha_4 T_t - \alpha_5 ST_t + \varepsilon_t \qquad \textbf{(3.10)}$$

where $\boldsymbol{\varepsilon_t}$ is a random error term while the other variables are as earlier defined.

Table 3.1 presents a summary of the variables used in Equation (3.9) and their corresponding names, as well as their measurement and expected signs.

**Table 3. 1: Description of variables in the empirical model and their hypothesized signs**

| Variable | Variable name | Description[7] | Measurement | Expected sign |
|---|---|---|---|---|
| $A_t$ | lnareac | Is the dependent variable representing the natural log of area planted with sugarcane | Hectares | |
| $A_{t-1}$ | $lnareac_{t-1}$ | Is the lagged dependent variable | Hectares | ( - ) |
| $p_t^y$ | lnrprice | Natural log of the price of cane deflated by CPI | Kes/Ton | ( + ) |
| $p_t^s$ | lnrmzprice | Natural log of the price of maize deflated by the CPI | Kes/90 kg bag | ( - ) |
| $y_t$ | lnyield | Natural log of the yield of sugarcane | Ton/Ha | ( + ) |
| ST | policy | A dummy variable representing change of company ownership from government to private | 1=privatized, 0 otherwise | ( + ) |
| T | T | Time trend variable representing technical change (state of technology) | Year | ( + ) |

Source: Author

In this study, the supply of sugarcane by outgrowers was hypothesized to be affected by price and non-price factors. Among the price factors is the cane price (own-price), prices of substitutes and prices of inputs. The first null hypothesis of the study stated that sugarcane price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run. Theoretically, an increase in the price of cane is likely to influence the cane supply in various ways (Nicholson and Snyder, 2008). Firstly, higher prices translate to higher farm incomes which may in turn enable them to purchase productivity-enhancing inputs in good time and apply them in proper quantities during the subsequent season. Secondly, higher prices may influence land allocation decisions by farmers towards producing more sugarcane in the following seasons in order to take advantage of the higher price.

[7] For all the variables that have a subscript *t*, the subscript represents the time period *t*.

The definition of the most competitive enterprise to sugarcane was based on the competition for available land for cultivation, which greatly influences farmers' land allocation decisions. The most competitive enterprise was determined by asking representative farmers through focus group discussions (FGDs) what crops take priority in terms of the proportion of area under different enterprises to total agricultural land cultivated in the respective districts within the study area. In Mumias cane growing zone, maize was found to be the most competitive enterprise. Data from the District and Provincial Annual Agricultural Reports confirmed this. The second null hypothesis stated that maize price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run.

Many factors affect yield: input application, amount and distribution of rainfall, soil characteristics, among others. Whatever the source of improvement in yield, such an increase is expected to improve the profitability of a commodity hence making it more attractive to the producers. As such, an increase in the yield of sugarcane was hypothesized to positively affect the area allocated to the crop in subsequent time periods.

Any policy shift is always intended to yield a positive growth in an industry. It was therefore hypothesized that the privatization of the sugar milling company should have addressed some of the constraints that the farmers faced, *ceteris peribus.* The expected sign for the time trend variable was hypothesized to be positive. This is because over the study period technological growth resulting from investment in infrastructural development and agricultural extension should have resulted in positive supply response.

### 3.2.2 Diagnostic tests

#### (a) Testing for stationarity

The first stage of the analysis involved the examination of the univariate properties of the time series data used in the model and estimating the order of integration. The time series variables were first independently tested for stationarity using the ADF (Komeh, 2012) and the Dickey Fuller Generalized Least Squares (DFGLS) tests. If the absolute value of the ADF test statistic was found to be smaller than the critical ADF value, then the null hypothesis that the series has unit root (non-stationary) could not be rejected. The non-stationary series were then tested to determine the order of integration using the ADF. This was operationalized by differencing the data and conducting the unit root test on the differenced data. This stepwise process was repeated until the series was found to be stationary and the number of differences conducted until the series became stationary represented the order of integration for the series.

Below is a summary of the ADF unit root test results for Mumias outgrowers. For all the series, with the exception of yield, the test included a trend.

**Table 3. 2: Augmented Dickey-Fuller Unit Root test results**

| Variable | Levels[8] (before differencing) | | First difference | | Order of Integration |
|---|---|---|---|---|---|
| | Test statistic | p-value | Test statistic | p-value | |
| LnAREAC | -3.60 | 0.03 | -6.16 | 0.00 | 1(0) |
| LnRPRICE | -2.39 | 0.39 | -5.41 | 0.00 | 1(1) |
| LnRMZPRICE | -2.85 | 0.18 | -6.43 | 0.00 | 1(1) |
| LnYIELD | -1.91 | 0.33 | -5.55 | 0.00 | 1(1) |

*[9]At levels, the critical values are -4.32, -3.57 and -3.22 for 1%, 5% and 10% significance levels respectively while at first difference they are -4.33, -3.58 and -3.23 for the respective levels of significance.

Source: Author

The ADF tested the null hypothesis that a certain series *X* has unit root. The rule is to reject the $H_0$ if the absolute value of the test statistic is greater than the critical value at the various significance levels or if the p-value is less than or equal to a specified significance level, i.e.,

[8] Levels data refers to the original data as reported, before any differences are taken.
[9] The critical values for the 1%, 5% and 10% significance levels are based on the number of observations and degrees of freedom. The levels series have 32 observations each while the first difference data has 31 observations. See Appendix Table A1.

one percent, five percent and 10 percent. If the $H_0$ is rejected it means that the series is stationary and hence has no unit root, i.e., it is I(0).

Table 3.2 shows that the $H_0$ was rejected at the 10 and 5 percent level of significance but could not be rejected at the 1 percent level for the natural log of area cultivated (lnAREAC). This meant that that at 1 percent level of significance the series was non-stationary. The series for the natural log of the real price of cane (lnRPRICE), natural log of the real price of maize (lnRMZPRICE) and natural log of yield (lnYIELD) were found to be non-stationary at the 1 percent level.

For confirmation purposes, an alternative test, the Dickey- Fuller Generalized Least Squares (DFGLS) test, which is considered a more powerful test (Cook, 2004) because it determines the presence of integration at different lag levels, was applied to all the series. The decision rule is to reject the $H_0$ (that the series has a unit root) when the computed statistic is greater than the critical values at the various levels of significance (StataCorp, 2009). The results showed that lnAREAC had a unit root when tested with three or more lags. The DFGLS test for the two price series (lnRPRICE and lnRMZPRICE) recommended a repeat of the test without the inclusion of any lags. Both were found to be non-stationary at zero lag order. The series lnYIELD was also found to be non-stationary using the DFGLS test (Appendix Table A2). All the series were therefore non-stationary at their levels.

After confirming the presence of unit root in all the levels series, the second step involved taking the first differences of the data and running the ADF unit root tests once again. In their first difference, all the series were found to be stationary. This led to the conclusion that all the series were I(1) since they were found to be stationary after being differenced once. The results for the unit root tests for the differenced series are presented in Appendix tables A3 and A4 for the ADF and DFGLS tests respectively, in the annex section.

Following the unit root tests for the data series on Mumias outgrowers, the first difference of all the series were used in the final model. Before estimating the VECM, however, the presence of co-integration among the variables had to be established.

### (b) Testing for co-integration

The second stage of the analysis involved testing the time series for co-integration. The Johansen co-integration test was used to test the time series data for co-integration. The Johansen co-integration test is based on the maximum likelihood (ML) estimation and two statistics; trace statistics and maximum Eigen values (Jonahsen, 1988). If the rank of the matrix is zero, then there is no co-integrating relationship. However, if it is greater than zero, then there are a number of co-integrating relationships equal to the maximum rank. The VECM specification takes into account the number of co-integrating relationships. The rank shows the number of long-run relationships that exist between the dependent variable and the explanatory variables. A rank of one means that only one linearly independent combination of the non-stationary variables will be stationary (Asari *et al.*, 2011).

Table 3.3 shows the Johansen co-integration test results for Mumias outgrowers data series.

**Table 3. 3: Johansen co-integration test results**

| Maximum rank | Eigen value | Trace statistic | 5% Critical value |
|---|---|---|---|
| r=0 | - | 74.70 | 68.52 |
| r=1 | 0.67 | 40.22* | 47.21 |
| r=2 | 0.50 | 18.78 | 29.68 |
| r=3 | 0.32 | 6.80 | 15.41 |
| r=4 | 0.16 | 1.23 | 3.76 |
| r=5 | 0.03 | - | - |

*denotes point at which to reject null hypothesis, i.e., where critical value exceeds trace statistic.
Source: Author's study

At a maximum rank of zero (r=0), the trace statistic (74.7) exceeds the critical value (68.52) and hence the null hypothesis of no co-integrating equations was rejected. However, at r=1, the trace statistic is lower than its critical value and hence the null hypothesis that there is at least one co-integrating equation could not be rejected. The conclusion was that there is at least one co-integrating equation among the series. As such the VECM for Mumias outgrowers was specified with the inclusion of one cointegrating equation.

**(c) Determination of optimal lags**

To determine the optimal number of lagged values of the explanatory variables to be included in the model, the *varsoc* test was applied in Stata. The test generates log-likelihood, likelihood ratios and values for three lag selection criteria: AIC, the HQIC and the SBIC. All three selection criteria suggested the inclusion of three lagged values per variable. Following results of the lag selection tests, the final VECM was specified using three lagged values per explanatory variable as well as for the lagged dependent variable.

**Table 3. 4: Lag selection test results**

| Lag | Log likelihood (LL) | Likelihood ratio (LR) | p | Final prediction error (FPE) | AIC (Akaike Information Criterion) | HQIC (Hannan and Quinn information criterion) | SBIC (Schwarz' Bayesian information criterion) |
|---|---|---|---|---|---|---|---|
| 0 | 18.60 | - | - | 0.023 | -0.938 | -0.864 | -0.702 |
| 1 | 18.67 | 0.14 | 0.707 | 0.025 | -0.874 | -0.785 | -0.590 |
| 2 | 18.75 | 0.17 | 0.684 | 0.026 | -0.811 | -0.707 | -0.480 |
| 3 | 23.95 | 10.40* | 0.001 | 0.020* | -1.100* | -0.982* | -0.723* |
| 4 | 24.19 | 0.48 | 0.487 | 0.021 | -1.048 | -0.915 | -0.624 |

The asterisks(*) show the optimal lag selection for the various selection criterion
Source: author’s study

### 3.2.3 VECM Specification

A VEC model was estimated following Equation (3.11) as follows.

$$\Delta lnAREAC_t = \beta_0 + \sum_{i=1}^{k-1} \beta_{i1}\Delta lnAREAC_{t-i} + \sum_{i=1}^{k-1} \beta_{i2}\Delta lnRPRICE_{t-i}$$

$$+ \sum_{i=1}^{k-1} \beta_{i3}\Delta lnRMZPRICE_{t-i} + \sum_{i=1}^{k-1} \beta_{i4}\,\Delta lnYIELD_{t-i} + \sum_{i=1}^{k-1} \beta_{i5}\,\Delta POLICY_t + \lambda\,[lnAREAC_t - \alpha_0$$

$$- \alpha_1 lnRPRICE_{t-1} - \alpha_2 lnRMZPRICE_{t-1} - \alpha_3 lnYIELD_{t-1} - \alpha_4 POLICY_t - \alpha_5 T_t]$$

$$+ \omega_t \qquad (\mathbf{3.11})$$

Where:

POLICY=1 when t>=2001; zero otherwise,

*K* is the number of lagged values for each explanatory variable , $\forall\ k = 2, ..., n$
$n=3$ and the other variables are as previously defined.

The specification of the VECM was based on the data properties of the time series variables when tested individually for stationarity and collectively for cointegration. The VECM generated parameters that show the responsiveness of sugarcane farmers to the price of cane, price of maize, and yield, both in the short and in the long-run, and an error correction mechanism which shows the speed of adjustment from short-run disequilibrium to long-run steady state.

Some additional tests to determine the reliability of the results were undertaken on the different VECM estimations. These include, the log likelihood ratio test for goodness of fit, the VECM stability test and the test for serial correlation of the error terms in the estimated model. Below is a presentation of the results obtained from these tests.

### 3.2.4 Post-estimation tests and procedures

#### a) Log likelihood ratio test

To test for the goodness of fit of the model a log-likelihood ratio was computed. Following Gujarati (2004), $LLR = 1 - (\frac{LLF_{ur}}{LLF_r})$ **(3.12)**

Where: LLR is the Log-likelihood ratio, $LLF_{ur}$ is the log-likelihood function for the model with all the variables while $LLF_r$ is the log-likelihood for the restricted regression that includes only the constant. $LLF_{ur}$ is equivalent to the residual sum of squares (RSS) while $LLF_r$ is equivalent to the total sum of squares (TSS) in a linear regression model.

#### b) Stability of the VECM

A test was undertaken to check the stability condition of the VECM estimates. The stability of a VECM refers to the ability of the system to revert back to the equilibrium after a shock (Asari *et al.*, 2011). The stability of linear dynamic systems can be determined from Eigen values (Woolf, undated). For a *K*-variable VECM with *r* co-integrating equations, the stability matrix will have *K*-*r* unit Eigen values. For stability, the moduli[10] of the remaining Eigen values should be strictly less than unity (StataCorp, 2009). Table 3.5 shows the stability condition of the VEC model for Mumias. The graphs of the Eigen values are presented as Figure A1 in the Annex Section as well.

[10] The moduli of a number is its non-negative or absolute value

**Table 3. 5: VECM stability test results**

| Eigen Values | | Modulus |
|---|---|---|
| 1 | | 1 |
| 1 | | 1 |
| 1 | | 1 |
| 1 | | 1 |
| -0.861 | | 0.861 |
| 0.465 | 0.714 | 0.852 |
| 0.465 | -0.714 | 0.852 |
| 0.211 | 0.793 | 0.820 |
| 0.211 | -0.793 | 0.820 |
| -0.504 | 0.499 | 0.710 |
| -0.504 | -0.499 | 0.710 |
| -0.283 | 0.420 | 0.506 |
| -0.283 | -0.420 | 0.506 |
| 0.402 | | 0.402 |
| 0.015 | | 0.015 |

Source: Author's study

The VECM had 5 variables, namely *lnAREAC, lnYIELD, lnRPRICE, lnRMZPRICE, ST* and was specified with one co-integrating equation (rank=1). The stability condition was therefore met in that there were 4 unit moduli and the rest of the modulus were all less than unity. As such, it was concluded that the estimates obtained from the VECM in this study was stable. This implies that the respective ECM terms are able to bring back the system to equilibrium after a shock (Asari *et al.*, 2011).

**c) Test for serial correlation**

The nature of time series data often results in correlated error terms as a result of inertia, the cobweb phenomenon, and data smoothening (Gujarati, 1999). With regard to model misspecification, under-specifying the number of lags in a VECM can significantly increase the finite-sample bias in the parameter estimates and lead to serial correlation (StataCorp, 2009). Autocorrelation implies that the least squares estimators are linear and unbiased but not efficient, while the variances are biased. As such, a Langrage multiplier test for autocorrelation on the residuals of the VECM was undertaken following Greene (2003). Table 3.6 shows the LM test results for the regression. The $H_0$ states that there is no autocorrelation at the respective lag levels for the lagged variables in the VECM. If the *p*-value for any of the lag levels is less than 0.1 then the $H_0$ is rejected at the respective significance level and the conclusion is that there is autocorrelation at that lag order (StataCorp, 2009).

**Table 3. 6: Autocorrelation test results**

| Lag | $Chi^2$ | $Prob>Chi^2$ |
|---|---|---|
| 1 | 13.46 | 0.97 |
| 2 | 31.85 | 0.16 |
| 3 | 16.30 | 0.91 |

Source: Author's study

Based on the results of the LM test, the $H_0$ of no serial correlation cannot be rejected even at the 10 percent level for all three lags in the model. This clearly indicates no evidence of autocorrelation in the residuals of the regression.

## 3.3 Data type and sources

### 3.3.1 Data sources

This study used secondary data. Secondary data were sourced from the Kenya Sugar Board annual publications as well as the Ministry of Agriculture's annual publications for maize prices. The study used company level data for outgrowers between 1980 and 2013 on sugarcane production aspects. This is the period for which a complete dataset on all the variables was available. The data series were on sugarcane tonnage supplied by outgrowers; area planted and harvested each year in hectares; cane yield (based on acreage harvested in the year, as opposed to acreage planted); and sugarcane producer prices. Cane supply was measured as the total cane (in tons) delivered by outgrowers to the milling factory while the cane price was the producer price paid per ton of cane delivered in Kenya shillings. To generate real prices for comparative purposes over the study period, the price data for sugarcane and maize were deflated using CPI with 2009 as the base year. The annual CPI data were obtained from the Kenya National Bureau of Statistics publications and website.

Reconnaissance visits to the Outgrowers Company were made at the initial stages of this study to provide background information and establish the nature of operations and relations between sugarcane farmers and the respective milling companies. Secondary data were then collected from the various sources outlined earlier.

### 3.3.2 Data limitations

A total of 34 observations were used in the analysis. It was difficult to obtain a longer time series for all the variables of interest in the study. While the standard errors and the probability of committing type I error decrease with an increase in the sample size (Gujarati, 1999), the computed statistics for hypothesis testing in this study followed the Z-distribution which follows the central limit theorem for large samples. Given that the number of observations $n>30$ there is reason to believe that the parameter estimates did not suffer small sample bias problem despite the sample size.

## 3.4 Data analysis

The data were captured in MS-Excel and analyzed using Stata and Eviews softwares. MS-Excel is good for handling relatively small datasets with ease, while Stata was used because it has more functionality and is robust for regression analysis than most analytical softwares. The analysis involved computation of descriptive statistics for the various series and running the MLE estimates of the VECM. The results were presented in charts and tables.

# CHAPTER 4: RESULTS AND DISCUSSION

## 4.1 Characteristics of sugarcane production systems among Mumias outgrowers

### 4.1.1 Yield and rainfall

Figure 4.1 shows trends in cane yield and rainfall for the period 1980 to 2013 for Mumias outgrowers. Plotting the sugarcane yield and mean annual amount of rainfall received within the cane growing zone showed that both variables were erratic over the period under study and without a clear trend overtime. While there were periods with wide gaps, the yield and average annual rainfall seemed to move together for the period between 1990 and 2009 indicating that, among other factors, rainfall is an important determinant of yield as expected. The time lag indicates that the yield in a particular time period is not only affected by rainfall received during the same period but also by that of the previous time periods. This is expected of sugarcane because it is a perennial crop.

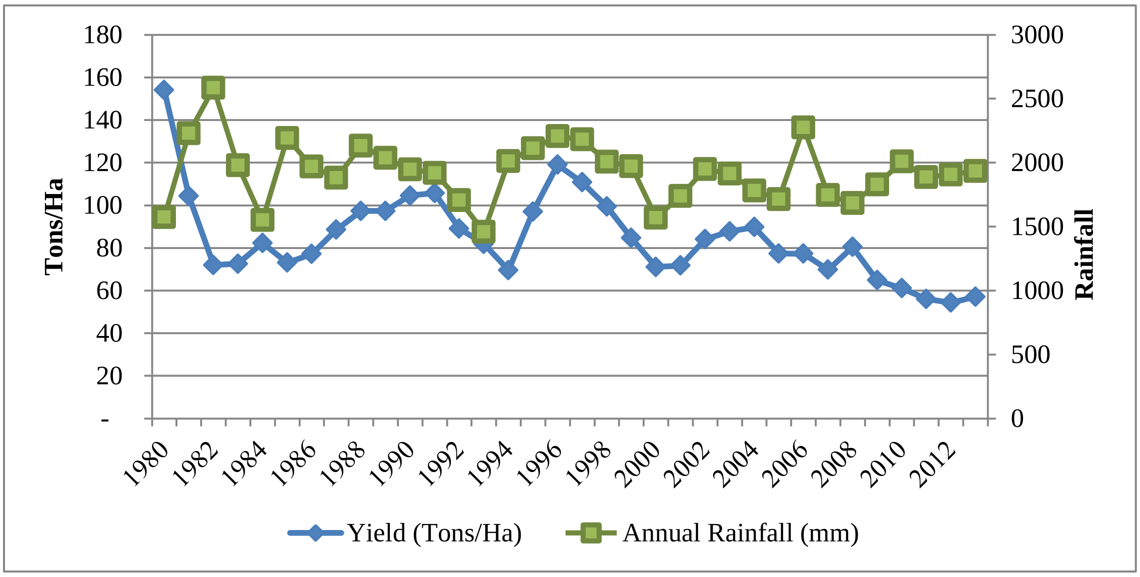



**Figure 4. 1: Sugarcane yield amongst outgrowers and rainfall received in the zone**
Source: KSB annual reports, various issues

Table 4.1 shows the summary statistics for the sugarcane yield and rainfall received in the Mumias cane growing zone. The yield achieved over the period ranged between 54 and 119 tons per hectare with a mean of 83 tons per hectare (se=2.9). The mean annual rainfall received over the period was 1939 mm per annum (se=40.3, range=1460-2585).

**Table 4. 1: Summary statistics for Mumias sugarcane yield and rainfall, 1981-2013**

| Statistic | Yield (Tons/hectare) | Annual Rainfall (mm) |
|---|---|---|
| Mean | 82.7 | 1939 |
| Standard error | 2.9 | 40.3 |
| Min | 54.3 | 1460 |
| Max | 119 | 2585 |
| n | 33 | 33 |

Source: Author's study

### 4.1.2 Prices

Figure 4.3 shows trends in real producer prices of sugar cane and maize in the Mumias cane growing zone. The prices for both cane and maize were rather constant in the 1980s. Starting 1990, the maize prices had a major spike which culminated in an all-time high in 1992, the only year when maize prices were higher than cane prices. This scenario could be explained by the fact that 1992 was an election year and the effect of pre-election uncertainty and post-election violence in the main maize growing areas resulted in a decline in production which translates to high prices due to excess demand over supply. After 1992 the maize prices were rather volatile and show a general decline which culminated in a major spike in 2008-2009, a period which was marked by excessively high food prices world over. The high maize prices in 2008-2009 were as a result of high production costs due to high fuel and fertilizer prices in the country.

Although the maize prices were volatile over the study period, the prices of cane were rather smooth. This could be attributed to the pricing strategy in which the milling company negotiates prices with the outgrower company as opposed to letting the market forces determine prices. Generally the prices of sugarcane were higher than those of maize in most of the period under study, but the gap was almost closed by the end of the study period (2013). The annual average maize prices have been on an upward trend since 2010 due to a deficit in production occasioned by the Maize Lethal Necrosis Disease (MNLD) in the main maize producing areas.

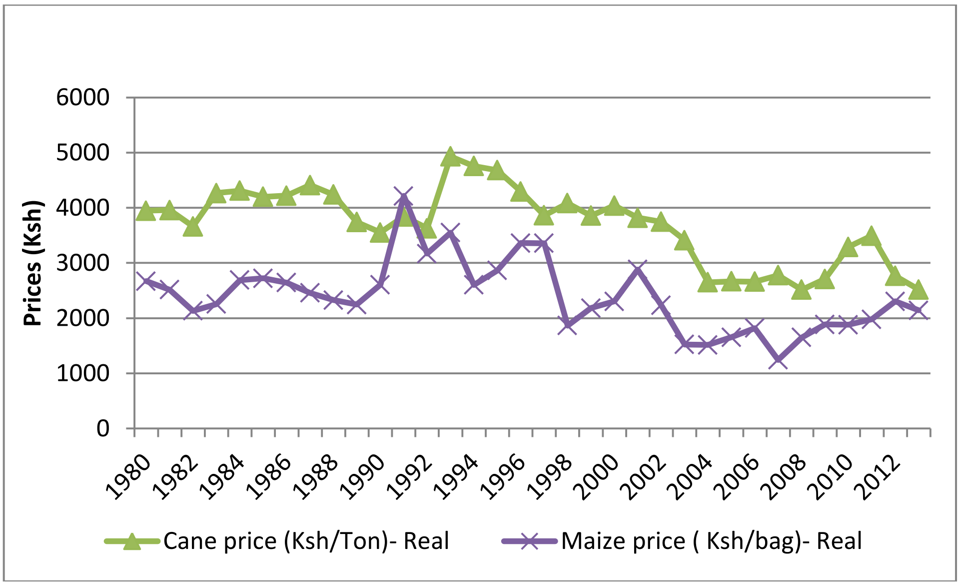



**Figure 4. 2: Real sugarcane and maize producer prices for the period 1981-2013 (2009=100)**

Source: KSB annual reports and MoA annual reports, various volumes

Table 4.2 presents a summary of the real prices received by outgrowers for sugarcane and the real maize producer prices within the Mumias cane growing zone over the period 1980 to 2013 with 2009 as the base year. The mean real price paid per ton of sugarcane was Ksh. 3683 (se=121) and ranged between Ksh. 2517 and Ksh. 4931 over the period. The mean real price of maize received was Ksh. 2387 (se=113; range=Ksh. 1246-4216).

**Table 4. 2: Summary statistics for real prices of sugarcane and maize for the period 1981-2013 (2009=100)**

| Statistic | Real cane price (Ksh/Ton) | Real maize price ( Ksh/90 kg bag) |
|---|---|---|
| Mean | 3683 | 2387 |
| Standard error | 121 | 113 |
| Min | 2,517 | 1246 |
| Max | 4,931 | 4,216 |
| n | 33 | 33 |

Source: Author's study

### 4.1.3 Discussion

Figure 4.1 shows that sugarcane yield and rainfall move together in Mumias cane growing zone. This is an indication that rainfall is a key determinant of yield. The range of sugarcane yield was 54-119 tons per hectare for Mumias outgrowers. These results compare well with Waswa *et al.* (2012) who found that the yield of sugarcane in Nzoia ranged between 72 and 75 tons per hectare for the period 2009-2011 while that of Mumias was 65-70 tons per hectare for the same period. Although these results are comparable with Reddy (2011) who found a mean productivity of 65 tons per hectare for sugarcane in India, there is a lot of yield potential that remains unutilized. Mandla (2012) found a mean productivity as high as 101 Tons/hectare in Swaziland. Data from FAOSAT shows that sugarcane yields in the Eastern Africa region are as high as 119 tons per hectare in Ethiopia and 98 Tons/hectare in Sudan. Improving the sugarcane yields would help bring the costs of production down and hence make Kenyan sugarcane more competitive.

The real producer prices of sugarcane were lower at the close of the study period in 2013 than they were in 1980. The same scenario was observed for maize prices. While maize prices were generally lower than cane prices, the gap was seen to converge towards the end of the study period. This is an indication that in terms of pricing, sugarcane lost more value relative to maize over the period. Although maize is a highly subsidized crop in Kenya, the decrease in the ratio of cane to maize prices is likely to make sugarcane farmers more attracted to maize production in their land allocation decisions.

The mean real cane price for the period under study was Ksh 3683/ton (range=2517-4931). Waswa *et al*. (2012) reported a range of Ksh 2150-4180/ton for Mumias for the period 2009-2012. The sugarcane prices in this study were closely comparable to Waswa *et al*. (2012) considering that the prices in this study were deflated using 2009 as the base year. The fact that real sugarcane prices lost more than real maize prices over the study implies that farmers in the sugarcane growing zones are likely to be attracted to maize production with the result that less land is allocated to sugarcane.

## 4.2 Supply response of sugarcane among outgrowers

### 4.2.1 VECM results

Table 4.3 shows the ML estimates for short-run and long-run supply response of Mumias sugarcane outgrowers. The table was extracted from Table A7 in the Annex section.

**Table 4. 3: Maximum Likelihood Estimates for short-run and long-run supply response**

| Variable | Coefficient | Standard Error | Test statistic (z) |
|---|---|---|---|
| **SHORT-RUN** | | | |
| $\Delta lnAREAC_{t-1}$ | 0.30 | 0.40 | 0.76 |
| $\Delta lnAREAC_{t-2}$ | 0.32 | 0.34 | 0.95 |
| $\Delta lnRPRICE_{t-1}$ | 0.64* | 0.37 | 1.74 |
| $\Delta lnRPRICE_{t-2}$ | 0.24 | 0.35 | 0.70 |
| $\Delta lnRMZPRICE_{t-1}$ | -0.15 | 0.18 | -0.86 |
| $\Delta lnRMZPRICE_{t-2}$ | -0.32 | 0.19 | -1.64 |
| $\Delta lnYIELD_{t-1}$ | -0.44 | 0.52 | -0.84 |
| $\Delta lnYIELD_{t-2}$ | 0.41 | 0.38 | 1.07 |
| $\Delta lnPRIVATE_{t-1}$ | -0.19 | 0.20 | -0.44 |
| $\Delta lnPRIVATE_{t-2}$ | 0.04 | 0.19 | 0.20 |
| CONSTANT | 0.03 | 0.03 | 0.81 |
| ECM | -1.11*** | 0.33 | -3.37 |
| **LONG-RUN** | | | |
| lnAREAC[11] | 1 | | |
| lnRPRICE | 0.72*** | 0.20 | 3.59 |
| lnRMZPRICE | -0.29*** | 0.09 | -3.37 |
| lnYIELD | 0.33*** | 0.12 | 2.69 |
| PRIVATE | -0.22** | 0.09 | -2.48 |
| TREND | -0.01*** | 0.002 | -4.04 |
| CONSTANT[12] | -14.90 | | |

Source: Author’s study

[11] In the formulation of a VECM the coefficient of the lagged dependent variable is always suppressed to 1 in the long-run parameters (StataCorp, 2009).
[12] The standard error and Z statistic for the constant term were blacked out to allow for the inclusion of a trend in the regression (StataCorp, 2009). The option would have been to specify a regression with a restricted constant term that has all the parameters but exclude the trend. The author acknowledges this limitation with the statistical package.

In the short-run, only the singly lagged price of the sugarcane (own price) significantly influenced sugarcane supply with the expected positive sign (p=0.08). This means that the area allocated to sugarcane is positively impacted by an increase in its price in the immediate preceding period. The coefficient was, however, less than unity (0.64) implying that the own-price elasticity of sugarcane supply was inelastic in the short-run. This means that when the prices of sugarcane increase the area allocated to the crop is likely to increase in the immediate succeeding period but that the increase in land allocation is relatively lower than the price change.

The error correction term (ECM) was statistically significant with the expected negative sign (p=0.001). This means that all the disequilibrium caused by any shock can be recovered within one time period (one year). With an ECM term of -1.11 sugarcane farmers in Mumias were able to recover from short-run flactuations in the area allocated to sugarcane and revert back to their long-run mean in a period of about 11 months i.e., 12/1.11.

In the long-run, all the explanatory variables were statistically significant at the 1 percent level with the exception of the privatization dummy variable which was significant at the 5 percent level (Table 4.3). The coefficient of the natural log of sugarcane price was 0.72 indicating that the own price elasticity of supply was inelastic in the long run. This implies that sugarcane farmers were not highly responsive to changes in the producer price of the commodity. The coefficient took the expected sign (positive) and although the magnitude of the coefficient was less than unity, the first null hypothesis which stated that sugarcane price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run was rejected since the parameter estimates were statistically significant for the singly lagged own price elasticity in the short-run (p=0.082) and p=0.000 for the long-run own price elasticity.

As expected, and in line with theory, the long-run own-price elasticity of sugarcane supply was higher (0.72) in the long-run than in the short-run (0.64). This may partly be explained by the fact that sugarcane is a perennial crop and hence adjustments in area allocated to sugarcane production due to higher prices require significant long-term capital investment. When prices fall, production continues at full capacity in order to spread the fixed costs, hence sugarcane supply tends to be own-price inelastic in the short-term (FAO, 2002). This implies that farmers have more flexibility to adjust their land allocation decisions in response to price changes in the long- as opposed to the short-run.

The coefficient on maize price took the expected negative sign but was inelastic (-0.29). This meant that a one percent decline in the price of maize would lead to a 0.29 percent increase in the land allocated to sugarcane in the successive period. The sign conforms to what was hypothesized in table 3.1. The second hypothesis of the study, (maize price has no effect on the supply of cane by outgrowers to Mumias Sugar Company both in the short- and long-run), was therefore rejected for the long-run but could not be rejected for the short-run. This implies that sugarcane outgrowers could switch to maize if maize prices increase and this would widen the sugar deficit gap in the country.

The yield elasticity of sugarcane supply was positive and elastic (0.33; p=0.007) in the long run. This was as expected a priori. With better yields, the profitability of the crop is expected to improve, *ceteris paribus*. As such, farmers are expected to allocate more land to sugarcane when the yield enhancing factors (higher rainfall, better use of inputs like fertilizers and improved seed material) lead to better yields hence more profitability.

The coefficient of the dummy variable for privatization was -0.22. This was contrary to expectation that privatization would affect sugarcane supply positively. The fourth null hypothesis which stated that there is no difference in the supply response of Mumias Sugar Company sugarcane farmers before and after privatization was therefore rejected.

The time trend was negative but statistically significant (p=0.000). This finding was contrary to the expectation as hypothesized in table 3.1. The magnitude of the coefficient was however very low (0.008) suggesting that there was very minimal technological change in the sugarcane sub-sector over the study period. The technological change however seems to have affected the supply response of sugarcane farmers in Mumias negatively. This could have been as a result of changing policy on extension service delivery from supply to demand driven approach and probably deteriorating road infrastructure.

To test for the goodness of fit of the model a log-likelihood ratio test was conducted. Regressing the variables in the model resulted in a RSS of 0.8 and TSS of 2.79. The log-likelihood ratio was therefore computed using the formula given in equation 3.11. The resultant log-likelihood ratio was 0.71. A log-likelihood ratio closer to one implies a better fit showing that the model fits the data well (Gujarati, 2004). The log-likelihood ratio for the model implies that 71 percent of the variations in the dependent variable, *lnAREAC*, were explained by the explanatory variables in the model. The RMSE was 0.16 and significant at 1 percent level. This shows that the model fitted the data well since a low RMSE shows a low variation between the sample and population estimators hence it implies more accuracy.

### 4.2.2 Discussion

Only the singly lagged own price of sugarcane was found to significantly affect the area allocated to the crop in the short-run with a coefficient of 0.64. This finding compares closely with Mubarik (1988) who estimated a short-run own price elasticity of sugarcane supply of 0.52 in Pakistan. Mythili (2006) found an output price elasticity of 0.26 for sugarcane in India.

In the long-run, the own-price elasticity of sugarcane supply was positive (0.72). A positive own-price elasticity of sugarcane supply shows that when prices increase farmers allocate more land to sugarcane. This finding implies that the farmers supplying the sugar milling factory were not very responsive to changes in the producer price of sugar cane. As such price support in sugarcane has little potential to increase supply. The results are comparable to Mubarik (1988) who found an own-price elasticity of sugarcane supply of 0.81 in Pakistan, while Piya (2002) found a long-run own price elasticity of sugarcane supply of 0.43. Harrington and Dubman (2008) found an implicit acreage response for sugarcane in the USA of 1.28 showing that sugarcane was a highly responsive crop to changes in its own price in the long-run. The difference in findings could be explained by the fact that supply response in developing countries tends to be less elastic than in developed countries due to infrastructural constraints and lack of complimentary policies and subsidies (Bingxin *et al.*, 2010).

The cross-price elasticity of sugarcane supply with respect to maize was negative and inelastic in the long-run (-0.29). This implies that an increase in the price of maize would lead to a decline in the land allocated to sugarcane in the long run as more land is allocated to maize since the maize price increase would make farmers deem maize a more attractive enterprise in comparison to sugarcane. This finding tallies with that of Schluter (1984) who found maize and sugarcane to be technical substitutes in Western Kenya.

The sugarcane yield was found to positively affect sugarcane supply in the long-run with an elasticity of 0.32. This implies that, in the long-run, a one percent increase in the yield of sugarcane would lead to a 0.32 percent increase in the area allocated to the crop in the following season. Muhhamad *et al.,* (2012) estimated a yield response of 0.21 for sugarcane in Pakistan while Jaforullah (1992) found a much higher sugarcane area response to own yield of 1.23 in Bangladesh.

In the MLE for Mumias, the coefficient of the time trend variable was -0.01 and highly significant (p=0.00). This is an indication that there have not been sufficient technical and technological growth in terms of investment in supportive infrastructure (roads and irrigation systems), education and extension service delivery as to lead to increased sugarcane supply. Additionally, an improvement in technology increases the marginal productivity of inputs like fertilizer hence a negative trend coefficient is not desirable. Mubaruk (1988) reported a trend coefficient of 0.023 which was attributed to research and infrastructural growth in Pakistan.

# CHAPTER 5: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Summary

The main focus of this study was to assess how responsive sugarcane farmers supplying Mumias Sugar Company are to changes in the producer price of sugarcane, price of a competing enterprise (maize) and to structural adjustments, specifically the change of company ownership from government control to private ownership. The main objectives were to characterize sugarcane production in Mumias and assess the supply response of the sugarcane farmers. Secondary data were collected from the MoA annual reports and the Kenya Sugar Board Annual publications. A vector error correction model (VECM) was fitted to the data and MLE estimates which indicated the partial responsiveness of the outgrowers to the various explanatory variables generated. The model included a dummy variable to represent the change of milling company ownership from a parastatal to public ownership in 2001 and a time trend variable to capture the effects of technological advancement. The analysis was conducted in STATA.

The study found that in the short-run, only the price of cane had a significant effect on the land area allocated to the crop (p=0.082). The coefficient was, however, less than unity, implying that the own-price elasticity of sugarcane supply was inelastic. In the long-run, all the variables in the model were found to significantly affect the area allocated to sugar cane. The own price and yield elasticities were positive, while the cross-price elasticity of maize to sugarcane supply was negative as hypothesized *a priori*. The time trend and privatization coefficients were, contrary to expectation, negative in the long-run. The error correction mechanism was greater that unity (-1.11) and significant indicating that the farmers were able adjust their production and revert back to the long term mean within a year after shock.

## 5.2 Conclusions

Given that only the own-price elasticity was significant in the short-run, it can be concluded that farmers in Mumias were not flexible in their land allocation decisions in the short-run. The fact that all the variables were significant in the long-run implies that the farmers needed time to adjust their production in response to changes in prices and other variables. This implies that sugarcane farming has an aspect of asset fixity which could be as a result of higher capital requirements when compared to other crops.

The study found that, although all the variables significantly affected the area response in sugarcane farming, none of them was elastic. This finding implies that the changes in prices and other variables in the producer environment were not strong enough to induce an elastic supply response. As such it can be concluded that sugarcane is not a highly responsive crop or that the sugarcane farmers in Mumias were constrained by factors that could not be empirically estimated within the scope of this study. The farmers could be facing a captive value chain governance structure where the oligopsonistic Mumias Sugar Company dictates when, how and how much land is to be allocated to sugarcane.

The positive sign of the own-price elasticity of sugarcane supply indicates that sugarcane farmers were responsive to changes in the prices paid per ton of cane delivered. Thus, farmers would increase the area allocated to sugarcane as a result of an increase in sugarcane prices in preceding periods. This suggests that a price incentive would be effective in increasing sugarcane deliveries by outgrowers in Mumias Sugar Company, *ceteris paribus*. However, issues of delayed payments, untimely harvesting, and faulty measurements should be urgently addressed.

The negative and significant cross price elasticity of maize to sugarcane supply implied that maize was indeed a competing enterprise to sugarcane in the outgrowers' land allocation decisions. This finding implies that the farmers were responsive to price signals and may consider maize a more attractive enterprise in comparison to sugarcane when maize prices rise. While this may not auger well for the self-sufficiency policy motive, it is a positive finding in that it suggests that the sugarcane farmers have a viable option towards the goal of attaining food security and poverty alleviation in the event that maize is actually a more profitable enterprise. The profitability of the two enterprises needs to be empirically tested before such a conclusion can be made.

The sugarcane outgrowers were found to respond positively to increases in yield in the long-run. While this can be viewed as a positive finding, increasing the land allocated to sugarcane in the face of land constraints may not be the optimal solution to shortages in supply of sugarcane. Instead of putting more land under cane, the farmers can opt for input use intensification which is a more effective measure towards higher farm profitability in the face of land constraints. This will also result in increased sugarcane supply.

The time trend elasticity was negative and significant in the long-run. This was an indication that there has not been much technical and technological growth in the sub-sector. Lack of growth in support services like education, extension and infrastructure impedes the marginal productivity of the variable inputs. The sugar cane farmers may therefore have been facing declining variable input productivity as a result.

The dummy for policy shift (privatization) was significant. This means that there indeed exists a difference between the parameter estimates between the periods before and after privatization of Mumias Sugar Company.  The fact that the coefficient was negative needs further investigation in order to understand why sugarcane farmers would respond negatively to privatization. Further, the problems of mismanagement in the Sugar milling company can be better analyzed under a political economy perspective than in a quantitative analytical framework. Such issues were beyond the scope of this study.

## 5.3 Policy recommendations

Based on the findings of this study the following recommendations were made:

1. Since the sugarcane farmers were found to respond positively to increases in the output price, strategies geared towards improving the profitability of the enterprise should be put in place as a long-term strategy to achieve self-sufficiency in sugar production in the country. Such strategies should not include price support, since it encourages inefficiencies, but should focus on the timeliness of both harvesting and payment to farmers for cane deliveries.
2. An increase in the price of maize was found to affect the supply of sugarcane negatively. In this regard, policies geared towards making sugarcane a more attractive enterprise so as to influence the farmer land allocation decisions positively are needed if the country is to attain self-sufficiency in sugar production. These include more timely payments for cane deliveries and advance payments so as to enable farmers acquire the needed inputs for subsequent seasons. The option would be to encourage the farmers to switch to the alternative enterprise if they deem it more profitable.
3. The privatization of government-owned sugar milling companies cannot be expressly recommended based on the findings of this study. If privatization is to be implemented as a long-term strategy, a more in-depth study that is of a qualitative nature and that captures the issues in the sub-sector from a political economy view point is needed.

4. The time trend coefficient was negative and highly significant. There is need to invest in rural infrastructure and appropriate technological improvements as a long-term strategy to improve sugarcane supply in the study areas. Such an investment would

lead to an improvement in the marginal productivity of other inputs like fertilizers and land and hence improve the profitability of sugar cane.

5. Sugarcane yield positively influenced the amount of land allocated to the crop in the long-run amongst outgrowers. Therefore, strategies to improve the use of yield-enhancing inputs like fertilizers are needed. Such strategies include making fertilizers available to the farmers at the right time, and the right price. Charging of interest on fertilizers supplied to cane farmers on credit is not efficient since on many occassions the cane overstays in the farms and the payments for cane deliveries are also delayed. Due to interest compounding this often leads to very high fertilizer prices and reduces the profitability of the enterprise. This scenario could also lead to low fertilizer application rates.

## 5.4 Recommendations for further research

1. The supply response of sugarcane outgrowers was found to have declined during the era of privatization. A further study to investigate whether there could be other issues that led to this scenario is recommended. Besides, it would be important to estimate the supply response of sugarcane farmers when a longer time series of the variable on privatization is available.

2. This study noted that the prices faced by maize and cane farmers are not efficient as they are controlled for the latter and subsidized for the former. While this study took the prices as given, it would be insightful to conduct a similar study using a price index that is more representative of perfect market conditions as an explanatory variable. When the results are compared it will be possible to gauge how the sugarcane farmers would respond to prices under a typical firm scenario.

3. Since the sugarcane farmers were found to respond to increases in maize prices, a study on the relative profitability of maize and sugarcane would help get a clear message on which of the two enterprises can benefit the farmers more. This kind of information will help sugarcane farmers make an informed decisions especially in the face of competition from cheap sugar imports from with COMESA once the one year extension of the COMESA sugar safegueards expires.

## REFERENCES


Alemu, Z.G., Oosthuizen, K., and van Schalkwyk, H.D., (2003). Grain-Supply Response in Ethiopia: An Error-Correction Approach. *Agrekon,* 42(4):391-392.

Asari, F.F., Baharuddin, N. S., Jusoh, N., Mohamad, Z., Shamsudin, N. and Jusoff, K., (2011). A Vector Error Correction Model (VECM) Approach in Explaining the Relationship between Interest Rate and Inflation towards Exchange Rate Volatility in Malaysia. *World Applied Sciences Journal*, 12: 49-56.

Askari, H. and Cummings, J.T., (1977). Estimating Agricultural Supply Response with the Nerlove Model: A Survey. *International Economic Review*, 18(2):257-292.

Banerjee, A., Dolado, J.J., and Mestre, R., (1998). Error Correction Mechanism Tests for Cointegration in a Single-Equation Framework. *Journal of Time Series Analysis*, 19(3): 267–283.

Bingxin, Y., Fengwei, L., and Liangzhi Y., (2010). Dynamic Agricultural Supply Response Under Economic Transformation: A Case Study of Henan Province. IFPRI Discussion Paper 00987, May 2010.

Buigut, S., (2011). A Fast-Track East African Community Monetary Union? Convergence Evidence from a Cointegration Analysis. School of Business Administration, American University in Dubai, UAE., pp 257.

Clark, J. S., and Klein, K.K., (1996). Nerlovian Area Response as an Error Correction Model: An Application to Western Canadian Agriculture. *Empirical Economics,* 21:501-512. Alberta Canada. http://link.springer.com/article/10.1007%2FBF01180698#page-1

Cook, S., (2004). Finite-Sample Properties of the GLS-Based Dickey-Fuller Test in the Presence of Breaks in Innovation Variance. Austrian Journal of Statistics, Volume 33 (3): 305–314.

Croser, J., and Anderson, K., (2010). Agricultural Distortions in Sub-Saharan Africa. Trade and Welfare Indicators, 1961 to 2004. The World Bank Development Research Group Trade and Integration Team June 2010, WPS 5344, pp 1-39.

Dickey, D. A., Fuller, W. A., (1979). Distribution of the Estimators for Autoregressive Time Series with Unit a Root. *Journal of the American Statistical Association*, 74 (366): 427-431.

FAO, (2013). Analysis of Incentives and Disincentives for Sugar in Kenya. Technical Notes Series, Monitoring African Food and Agricutural Policies, FAO, Rome.

FAO, (2002). Agricultural commodities: Profiles and relevant WTO negotiating issues. http://www.fao.org/DOCREP/006/Y4343E/Y4343E00.HTM Accessed April 2014

Gicheru, S., Waiyaki, N., Omiti, J., (2007). Technical Efficiency of Kenya's Sugar Factories: an Agenda for Enhancing Competitiveness. Kenya Institute for Public Policy and Research Analysis. Nairobi, pp 37-38.

Greene ,W. H., (2003). *Econometric Analysis*. 5th Edition. Prentice hall. New Jersey.

Gujarati, D., (2004). *Basic Econometrics*, 4th Edition, 2004. McGraw-Hill International. Singapore.

Gujarati, D., (1999). *Essentials of Econometrics*, 2nd Edition, 1999. McGraw-Hill International. Singapore.

Hansen, P. R., (2004). Granger's representation theorem: A closed-form expression for *I*(1) processes .*Econometrics Journal*, 8:23–38.

Harrington, D. H., and Dubman R., (2008). Equilibrium Displacement Mathematical Programming models: Methodology and a model for the U.S Agricultural Sector. USDA Technical Bulletin Number 1918.

Hassler, U., and Wolters, JÄurgen., (2006). Autoregressive Distributed Lag Models and Cointegration. *Allgemeines Statistisches Archive*, 0:1-14.

Haughton, J., (1985). Farm Price Responsiveness and the Choice of a Functional Form: an Application to Rice Cultivation in West Malaysia. *Journal of Development Economics,* 24:203-223.

Hendry, D. F. and Juselius, K., (2000). Explaining Cointegration Analysis: Part I. *Energy Journal*, 21:1–42.

Hill, R. C., Griffiths W. E., and Lim G. C., (2011). *Principles of Econometrics*, 4th edition 2011. Willy and Sons Inc. New Jersey.

Hornsby, C., (2012). *Kenya: A History since Independence*. I. B Tauris & Co., London.

Jaforullah, M., (1992). Sugarcane area response in the mill zones of Bangladesh 1947/48-1971/72. *The Developing Economies*, 3(1992):236-258.

Jayne, T.S., Myers R.J. and Nyoro, J., (2005). Effects of Government Maize Marketing and Trade Policies on Maize Market Prices in Kenya. Tegemeo Institute of Agricultural Policy and Development, WPS 15/2005, pp 1-39.

Johansen S., (1987). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*. North-Holland, University of Copenhagen, Denmark. 12 (1988): 231-254.

Kapuya, T., (2010). Modelling the Impact of the 'Fast-track' Land Reform Policy in Zimbabwe's Maize Sector. Unpublished Msc. Thesis. University of Pretoria, South Africa.

Kenya Sugar Board, (2010). Kenya Sugar Industry Strategic Plan 2010-2014. Government Printer, Nairobi.

Kiiru, M.W., (1995). Determination of the Price Elasticity of Sunflower Supply in Kenya. Unpublished Msc. Thesis, University of Nairobi.

Kipruto, K.B., (2010). Effect of Trade Protectionism on the Efficiency of the Sugar Industry in Kenya. In Proceedings of the 12th KARI Biennial Scientific Conference, Nairobi Kenya, pp 1279-1283.

Komeh, T. F., (2012). The effects of private investment on growth in sub-Saharan Africa between 1990-2008. Unpublished Masters Thesis. Jönköping International Business School.

Kuwomu, J.K., Maanikuu P. M. , and Yaw B. O., (2011). Supply Response of Rice in Ghana: A Co-integration Analysis. Department of Agricultural Economics and Agribusiness, University of Ghana, Legon-Accra, Ghana. *Journal of Economics and Sustainable Development*, 2(6):1-14.

Lütkepohl, H., (2005). Structural Vector Autoregressive Analysis for Cointegrated Variables. Department of Economics European University Institute, EUI Working Paper ECO No. 2005 /2, San Domenico, Italy.

Lütkepohl, H. and Kratzig M., (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

Maddala, G. S., (2002). *Introduction to Econometrics*, 3rd Edition. Wiley India Pvt. Ltd. New Delhi.

Mandere, N.M., (2003). Towards Sustainable Agriculture: Sugar Beet Production in Kenya as a Potential Viable Solution to the Current Sugar Deficit. Unpublished Msc. Thesis, University of Lund, Sweden.

Mandla, B. D., and Masuku, M. B., (2012). Productivity of Smallholder Sugarcane Farmers in Swaziland: The Case of Komati Downstream Development Programme (KDDP) Farmers' Associations, 2005-2011. Canadian Centre of Science and Education. *Environment and Natural Resources Research*, 2 (4):1-7.

Mbogoh, S.G., (1980). An Economic Analysis of Kenya's Sugar Industry with Special Reference to the Self-sufficiency Production Policy. Unpublished Phd Thesis, University of Alberta, Canada.

Mesike, C.S., Okoh R.N. and Inoni, O.E., (2010). Supply Response of Rubber Farmers in Nigeria: An Application of Vector Error Correction Model. *Agricultural Journal*, 5(3):146-150.

Mohammad, S., Javed, M. S., Ahmad, B., and Mushtaq, K., (2007). Price and Non-price Factors Affecting Acreage Response of Wheat in Different Agro-ecological Zones in Punjab: A Co-integration Analysis. *Pak. J. Agri. Sci.,* 44(2):370-377.

Mose, L.O., (2007). Who Gains, Who loses? The Impacts of Market Liberalization on Rural Households in Northwestern Kenya. Unpublished PhD Thesis, Wageningen University, the Netherlands.

Mubarik A., (1988). Supply Response of Major Crops in Pakistan. A Simultaneous Equation Approach. Chemonics International Consulting Division, Islamaband.

Muchapondwa, E., (2009). Supply Response of Zimbabwean Agriculture: 1970–1999. *African Journal of Agricultural and Resource Economics*, 3(1):28-42.

Mulwa, R., Nuppenau, E., Emrouznejad, A., (2005). Productivity Growth in Smallholder Sugarcane Farming in Kenya: A Malmquist TFP Decomposition. A paper presented in the Tropentag, Stuttgart Germany, October, 2005.

Mythili, G., (2008). Acreage and Yield Response for Major Crops in the Pre- and Post-Reform Periods in India: A Dynamic Panel Data Approach. Report Prepared for IGIDR – ERS/ USDA Project: Agricultural Markets and Policy. *Indira Gandhi Institute of Development Research*, Mumbai, India.

Mythili, G., (2006). Supply Response of Indian Farmers: Pre and Post Reforms. *Indira Gandhi Institute of Development Research*, Mumbai, India.

Nerlove, M., (1956). Estimates of the Elasticities of Supply of Selected Agricultural Commodities. *Journal of Farm Economics*, 38:496-509.

Nicholson, W., Snyder, C., (2008). *Microeconomic Theory Basic Principles and Extensions,* Tenth Edition. Thomson Inc., Ohio.

Nkang, N., Ndifoni, H., Edet, E., (2007). Maize Supply Response to Changes in Real Prices: A Vector Error Correction Approach. *Agricultural Journal,* 2(3):419-425.

Nyongesa, D.P., (2003). Review of Cane Payment Systems: A Case Study of the Kenya Sugar Industry. *In Proceedings of the 11th Biennial Congress, Kenya Society of Sugarcane Technologists*, Mumias, Kenya, pp. 113-115.

Obayelu, A.E., and Salau, S.A., (2010). Agricultural Response to Prices and Exchange Rate in Nigeria: Application of Co-integration and Vector Error Correction Model. *Agri Sci,* 1(2):73-81.

Odenya, J.O., Ochia , C.O., Korir, C., . Otieno, V., and Bor, G.K., (2010). Adoption of Improved Sugarcane Varieties in Nyando Sugarcane Zone, Kenya. KESREF.

Odhiambo, G.O., (2008). Partial Equilibrium Analysis of the Implications of Import Tariff Elimination on the Sugar Industry in Kenya. Unpublished Msc. Thesis, Department of Agricultural Economics, University of Nairobi.

Olubode-Awosola, O.O., Oyewumi, O.A., and Jooste, A., (2006). Vector Error Correction Modelling of Nigerian Agricultural Supply Response. *Agrekon,* 45(4):421-436.

Olwande, J.O., Ngigi, M., Nguyo, W., (2009). Supply Responsiveness of Maize Farmers in Kenya: A Farm-Level Analysis. Contributed paper presented at the International Association of Agricultural Economists (IAAAE) 2009 Conference, held in Beijing China.

Ozkan, B. , Ceylan, R. F., Kizilay, H., (2011). Supply Response for Wheat in Turkey: A Vector Error Correction Approach. Faculty of Agriculture, Department of Agricultural Economics, Akdeniz University, Antalya Turkey.

Pfumayaramba, T., (2011). Analysis of Flue-cured Tobacco Supply Elasticity in Zimbabwe 1980-2010: An Error Correction Model Approach. Paper Prepared for Presentation at the 55th Annual AARES National Conference Melbourne, Victoria.

Piya, S., (2002). The Effect of Price and Non-price Factors on Agricultural Production in Nepal: A Cointegration Analysis. Department of Agricultural and Resource Economics, The University of Tokyo.

Reddy A. A., (2011). Sugar and Cane Pricing and Regulation in India. International Crops Research Institute for Semi-arid Tropics (ICRISAT). Andhra Pradesh India. *International Sugar Journal*, 113(1352):548-556.

Republic of Kenya, (2010). Economic Review of Agriculture. Central Planning and Project Monitoring Unit, Ministry of Agriculture. Government Printer, Nairobi.

Republic of Kenya, (2011). Economic Survey. Kenya National Bureau of Statistics, Government Printer, Nairobi.

Ruffer, T., and Vergano, P., (2002). *An Agricultural Safeguard Mechanism for Developing Countries*. Oxford Policy Management and O’Conner and Company, Oxford, UK.

Sadoulet, E., and de Janvry, A., (1995). *Quantitative Development Policy Analysis*. Johns Hopkins University Press , Baltimore, Maryland, pp. 61-111.

Schluter M., (1984). Constraints on Kenya's Food and beverage exports. Research report 44 IFPRI.

Sjo, B., (2008). Testing for Unit Roots and Cointegration. Unpublished lecture notes, pp 10-12. file:///H:/MORE%20THESIS%20PAPERS/Cointegration%20and%20unit%20root%20test%20equations.pdf Accessed November 2013.

StataCorp, (2009). Stata Time Series Reference Manual Release 11. *Stata Press*, 4905 Lakeway Drive, College Station, Texas 77845.

Thiele, R., (2002). Price Incentives, Non-Price Factors, and Agricultural Production in Sub-Saharan Africa: A Cointegration Analysis. Kiel Institute for World Economics. Working paper no. 1123. Duesternbrooker Weg 120 24105 Kiel, Germany.

Ullah, R., Ali, S., Safi, Q. S., Shah, J. and Khan, K. H., (2012). Supply Response Analysis of Wheat Growers in District Peshawar: Pakistan. *Agr. Food Sci*, 2(1):33-38.

USDA, (2014). Kenya Sugar Annual Report. Global Agricultural Information Network, Nairobi. http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Sugar%20Annual_Nairobi_Kenya_4-15-2014.pdf. Accessed April 2015.

USDA, (2011). Kenya Sugar Annual Report. Global Agricultural Information Network, Nairobi. http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Sugar%20Annual_Nairobi_Kenya_4-14-2011.pdf. Accessed January 2012.

Varian, H. R. (1992). *Microeconomic Analysis*. 3rd edition, New York.

Woolf, P., (undated). Dynamical Systems Analysis II: Evaluating Stability, Eigenvalues. Unpublished lecture notes. University of Michigan. https://controls.engin.umich.edu/wiki/images/3/34/Lecture.15.pdf. Accessed April 2014.

Zulfiqar, M., Chishti, A.F., Khan, M., Waheed, A., Wasiullah, Shah, M., Zakir, M., Arifullah, S.A., Khan, D., Sarfaraz, N., and Amin, H., (2011). Sugar and Cane Pricing and Regulation in India. *Sarhad J. Agric*, 27(2):285-290.

# APPENDICES

## A.1: AUGMENTED DICKEY-FULLER TESTS

### Table A 1: Augmented Dickey-fuller tests- data in levels

```
. dfuller lnareac, trend

Dickey-Fuller test for unit root                   Number of obs   =        32

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -3.596            -4.316            -3.572            -3.223
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0302

. dfuller lnrprice, trend

Dickey-Fuller test for unit root                   Number of obs   =        32

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.385            -4.316            -3.572            -3.223
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.3874

. dfuller lnrmzprice, trend

Dickey-Fuller test for unit root                   Number of obs   =        32

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.849            -4.316            -3.572            -3.223
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.1796

. dfuller lnyield

Dickey-Fuller test for unit root                   Number of obs   =        32

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.907            -3.702            -2.980            -2.622
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.3290
```

## Table A 2: Dickey-Fuller GLS tests

```
. dfgls lnareac

DF-GLS for lnareac                                       Number of obs =    23
Maxlag = 9 chosen by Schwert criterion

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    9            -1.362           -3.770            -2.853            -2.405
    8            -1.010           -3.770            -2.808            -2.407
    7            -1.344           -3.770            -2.820            -2.453
    6            -1.323           -3.770            -2.877            -2.531
    5            -1.581           -3.770            -2.967            -2.632
    4            -2.195           -3.770            -3.078            -2.747
    3            -2.604           -3.770            -3.197            -2.864
    2            -4.010           -3.770            -3.313            -2.974
    1            -3.889           -3.770            -3.414            -3.067

Opt Lag (Ng-Perron seq t) =  1 with RMSE  .1604522
Min SC   = -3.386867 at lag  1 with RMSE  .1604522
Min MAIC = -1.729339 at lag  8 with RMSE  .1462672

. dfgls lnrprice

DF-GLS for lnrprice                                      Number of obs =    23
Maxlag = 9 chosen by Schwert criterion

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    9            -1.494           -3.770            -2.853            -2.405
    8            -1.450           -3.770            -2.808            -2.407
    7            -1.134           -3.770            -2.820            -2.453
    6            -1.017           -3.770            -2.877            -2.531
    5            -1.530           -3.770            -2.967            -2.632
    4            -1.438           -3.770            -3.078            -2.747
    3            -1.606           -3.770            -3.197            -2.864
    2            -2.014           -3.770            -3.313            -2.974
    1            -2.077           -3.770            -3.414            -3.067

Opt Lag (Ng-Perron seq t) = 0 [use maxlag(0)]
Min SC   =  -4.17411 at lag  1 with RMSE  .1082425
Min MAIC = -3.876215 at lag  1 with RMSE  .1082425

. dfgls lnrprice,  maxlag(0)

DF-GLS for lnrprice                                      Number of obs =    32

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    0            -2.374           -3.770            -3.352            -3.029

. dfgls lnrmzprice

DF-GLS for lnrmzprice                                    Number of obs =    23
Maxlag = 9 chosen by Schwert criterion

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    9            -2.200           -3.770            -2.853            -2.405
    8            -1.900           -3.770            -2.808            -2.407
    7            -1.978           -3.770            -2.820            -2.453
    6            -2.030           -3.770            -2.877            -2.531
    5            -1.588           -3.770            -2.967            -2.632
    4            -1.354           -3.770            -3.078            -2.747
    3            -1.225           -3.770            -3.197            -2.864
    2            -1.824           -3.770            -3.313            -2.974
    1            -2.160           -3.770            -3.414            -3.067

Opt Lag (Ng-Perron seq t) = 0 [use maxlag(0)]
Min SC   = -2.784527 at lag  1 with RMSE  .2168414
Min MAIC = -2.586333 at lag  3 with RMSE  .2077565

. dfgls lnrmzprice, maxlag(0)

DF-GLS for lnrmzprice                                    Number of obs =    32

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    0            -2.903           -3.770            -3.352            -3.029

. dfgls lnyield

DF-GLS for lnyield                                       Number of obs =    23
Maxlag = 9 chosen by Schwert criterion

               DF-GLS tau      1% Critical       5% Critical      10% Critical
  [lags]     Test Statistic        Value             Value             Value
------------------------------------------------------------------------------
    9            -0.844           -3.770            -2.853            -2.405
    8            -0.852           -3.770            -2.808            -2.407
    7            -0.881           -3.770            -2.820            -2.453
    6            -0.857           -3.770            -2.877            -2.531
    5            -0.942           -3.770            -2.967            -2.632
    4            -1.118           -3.770            -3.078            -2.747
    3            -1.610           -3.770            -3.197            -2.864
    2            -2.786           -3.770            -3.313            -2.974
    1            -2.822           -3.770            -3.414            -3.067

Opt Lag (Ng-Perron seq t) =  1 with RMSE   .114006
Min SC   = -4.070357 at lag  1 with RMSE   .114006
Min MAIC = -3.695775 at lag  5 with RMSE  .1055074
```

**Table A 3: Augmented Dickey-fuller tests- series in first difference**

```
. dfuller d.lnareac, trend

Dickey-Fuller test for unit root                   Number of obs   =        31

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -6.155            -4.325            -3.576            -3.226
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

. dfuller d.lnrprice, trend

Dickey-Fuller test for unit root                   Number of obs   =        31

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -5.412            -4.325            -3.576            -3.226
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

. dfuller d.lnrmzprice, trend

Dickey-Fuller test for unit root                   Number of obs   =        31

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -6.430            -4.325            -3.576            -3.226
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

. dfuller d.lnyield

Dickey-Fuller test for unit root                   Number of obs   =        31

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -5.549            -3.709            -2.983            -2.623
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

## Table A 4: Dickey-Fuller GLS tests- series in first difference

. dfgls d.lnareac

DF-GLS for D.lnareac Number of obs = 22
Maxlag = 9 chosen by Schwert criterion

| [lags] | DF-GLS tau Test Statistic | 1% Critical Value | 5% Critical Value | 10% Critical Value |
|---|---|---|---|---|
| 9 | -2.065 | -3.770 | -2.909 | -2.432 |
| 8 | -2.011 | -3.770 | -2.835 | -2.414 |
| 7 | -2.863 | -3.770 | -2.829 | -2.447 |
| 6 | -2.784 | -3.770 | -2.876 | -2.519 |
| 5 | -3.606 | -3.770 | -2.962 | -2.620 |
| 4 | -4.322 | -3.770 | -3.075 | -2.738 |
| 3 | -4.219 | -3.770 | -3.199 | -2.861 |
| 2 | -5.033 | -3.770 | -3.322 | -2.977 |
| 1 | -3.966 | -3.770 | -3.428 | -3.076 |

Opt Lag (Ng-Perron seq t) = 2 with RMSE .1802012
Min SC = -3.005857 at lag 2 with RMSE .1802012
Min MAIC = .2257974 at lag 1 with RMSE .2086102

. dfgls d.lnrprice

DF-GLS for D.lnrprice Number of obs = 22
Maxlag = 9 chosen by Schwert criterion

| [lags] | DF-GLS tau Test Statistic | 1% Critical Value | 5% Critical Value | 10% Critical Value |
|---|---|---|---|---|
| 9 | -1.185 | -3.770 | -2.909 | -2.432 |
| 8 | -1.440 | -3.770 | -2.835 | -2.414 |
| 7 | -1.534 | -3.770 | -2.829 | -2.447 |
| 6 | -2.162 | -3.770 | -2.876 | -2.519 |
| 5 | -2.913 | -3.770 | -2.962 | -2.620 |
| 4 | -2.346 | -3.770 | -3.075 | -2.738 |
| 3 | -2.960 | -3.770 | -3.199 | -2.861 |
| 2 | -3.213 | -3.770 | -3.322 | -2.977 |
| 1 | -3.163 | -3.770 | -3.428 | -3.076 |

Opt Lag (Ng-Perron seq t) = 0 [use maxlag(0)]
Min SC = -3.971395 at lag 1 with RMSE .1192898
Min MAIC = -1.949385 at lag 1 with RMSE .1192898

. dfgls d.lnrprice, maxlag(0)

DF-GLS for D.lnrprice Number of obs = 31

| [lags] | DF-GLS tau Test Statistic | 1% Critical Value | 5% Critical Value | 10% Critical Value |
|---|---|---|---|---|
| 0 | -5.482 | -3.770 | -3.365 | -3.039 |

. dfgls d.lnrmzprice

DF-GLS for D.lnrmzprice Number of obs = 22
Maxlag = 9 chosen by Schwert criterion

| [lags] | DF-GLS tau Test Statistic | 1% Critical Value | 5% Critical Value | 10% Critical Value |
|---|---|---|---|---|
| 9 | -2.043 | -3.770 | -2.909 | -2.432 |
| 8 | -1.881 | -3.770 | -2.835 | -2.414 |
| 7 | -2.202 | -3.770 | -2.829 | -2.447 |
| 6 | -1.997 | -3.770 | -2.876 | -2.519 |
| 5 | -1.743 | -3.770 | -2.962 | -2.620 |
| 4 | -2.480 | -3.770 | -3.075 | -2.738 |
| 3 | -3.542 | -3.770 | -3.199 | -2.861 |
| 2 | -5.629 | -3.770 | -3.322 | -2.977 |
| 1 | -5.270 | -3.770 | -3.428 | -3.076 |

Opt Lag (Ng-Perron seq t) = 2 with RMSE .1859061
Min SC = -2.943522 at lag 2 with RMSE .1859061
Min MAIC = 3.945269 at lag 1 with RMSE .2073972

. dfgls d.lnyield

DF-GLS for D.lnyield Number of obs = 22
Maxlag = 9 chosen by Schwert criterion

| [lags] | DF-GLS tau Test Statistic | 1% Critical Value | 5% Critical Value | 10% Critical Value |
|---|---|---|---|---|
| 9 | -1.867 | -3.770 | -2.909 | -2.432 |
| 8 | -1.945 | -3.770 | -2.835 | -2.414 |
| 7 | -1.977 | -3.770 | -2.829 | -2.447 |
| 6 | -2.027 | -3.770 | -2.876 | -2.519 |
| 5 | -2.314 | -3.770 | -2.962 | -2.620 |
| 4 | -2.676 | -3.770 | -3.075 | -2.738 |
| 3 | -3.440 | -3.770 | -3.199 | -2.861 |
| 2 | -4.066 | -3.770 | -3.322 | -2.977 |
| 1 | -3.303 | -3.770 | -3.428 | -3.076 |

Opt Lag (Ng-Perron seq t) = 2 with RMSE .1254703
Min SC = -3.729867 at lag 2 with RMSE .1254703
Min MAIC = -1.968867 at lag 1 with RMSE .1382655

## A.2: JOHANSEN COINTEGRATION AND LAG SELECTION TEST RESULTS

### Table A 5: Johansen co-integration test

```
. vecrank lnareac lnrprice lnrmzprice lnyield private

                       Johansen tests for cointegration
Trend: constant                                        Number of obs =      31
Sample:  1983 - 2013                                            Lags =       2
-------------------------------------------------------------------------------
                                                         5%
maximum                                     trace    critical
  rank    parms       LL       eigenvalue  statistic   value
    0      30     96.285911         .       74.6954    68.52
    1      39     113.52433     0.67115     40.2185*   47.21
    2      46     124.24216     0.49916     18.7829    29.68
    3      51     130.23203     0.32053      6.8031    15.41
    4      54     133.01836     0.16453      1.2305     3.76
    5      55      133.6336     0.03892
-------------------------------------------------------------------------------
```

### Table A 6: Lag selection test

```
. varsoc lnareac, exog (lnrprice lnrmzprice lnyield private)

   Selection-order criteria
   Sample:  1985 - 2013                         Number of obs      =        29
```

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | 18.5996 | | | | .022999 | -.937905 | -.864074 | -.702165 |
| 1 | 18.6703 | .14139 | 1 | 0.707 | .024585 | -.873815 | -.785218 | -.590927 |
| 2 | 18.7529 | .16518 | 1 | 0.684 | .026286 | -.810546 | -.707182 | -.480509 |
| 3 | 23.9537 | 10.402* | 1 | 0.001 | .019773* | -1.10025* | -.982124* | -.723068* |
| 4 | 24.1948 | .48216 | 1 | 0.487 | .020971 | -1.04791 | -.915018 | -.623581 |

```
   Endogenous:  lnareac
    Exogenous:  lnrprice lnrmzprice lnyield private  _cons
```

The asterisks (*) show the optimal lag for the various selection criterion.

## A.4: VECM RESULTS

### Table A 7: VECM results for Mumias

```
. vec lnareac lnrprice lnrmzprice lnyield private, rank(1) lags(3) trend(rtrend)

Vector error-correction model

Sample:  1984 - 2013                              No. of obs      =         30
                                                  AIC             = -5.691981
Log likelihood =  150.3797                        HQIC            = -4.720762
Det(Sigma_ml)  =  3.05e-11                        SBIC            = -2.656053

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
D_lnareac            12     .162129   0.6549   32.25991   0.0013
D_lnrprice           12     .107403   0.4113   11.87478   0.4558
D_lnrmzprice         12      .20033   0.4987    16.9119   0.1529
D_lnyield            12       .0892   0.7107   41.76766   0.0000
D_private            12     .167741   0.4935   16.56588   0.1667
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
D_lnareac    |
         _ce1 |
          L1. |  -1.110626   .3294912    -3.37   0.001    -1.756417   -.4648356
             |
     lnareac |
          LD. |   .3039852   .4008388     0.76   0.448    -.4816444    1.089615
         L2D. |   .3242227   .3406993     0.95   0.341    -.3435356     .991981
             |
    lnrprice |
          LD. |    .643152   .3695923     1.74   0.082    -.0812356     1.36754
         L2D. |   .2425926   .3460938     0.70   0.483    -.4357389     .920924
             |
  lnrmzprice |
          LD. |  -.1530952   .1771658    -0.86   0.388    -.5003338    .1941434
         L2D. |  -.3195305   .1942625    -1.64   0.100     -.700278    .0612171
             |
     lnyield |
          LD. |  -.4373989   .5192501    -0.84   0.400     -1.45511    .5803126
         L2D. |    .411473   .3842569     1.07   0.284    -.3416567    1.164603
             |
     private |
          LD. |  -.0897445   .2022094    -0.44   0.657    -.4860678    .3065787
         L2D. |   .0384877   .1946826     0.20   0.843    -.3430831    .4200585
             |
        _cons |   .0264488    .032553     0.81   0.417     -.037354    .0902516
------------------------------------------------------------------------------

Cointegrating equations

Equation           Parms    chi2     P>chi2
-------------------------------------------
_ce1                   4   170.1848   0.0000
-------------------------------------------

Identification:  beta is exactly identified

                 Johansen normalization restriction imposed
------------------------------------------------------------------------------
        beta |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_ce1         |
     lnareac |          1          .        .       .            .           .
    lnrprice |   .7196666   .2002837     3.59   0.000     .3271178    1.112215
  lnrmzprice |  -.2898354   .0860401    -3.37   0.001    -.4584709   -.1211999
     lnyield |   .3294077     .12226     2.69   0.007     .0897825    .5690329
     private |  -.2151788   .0866188    -2.48   0.013    -.3849484   -.0454092
       _trend |  -.0081443   .0020163    -4.04   0.000    -.0120961   -.0041925
        _cons |  -14.8991          .        .       .            .           .
------------------------------------------------------------------------------
```

## A.5: VECM STABILITY TESTS

### Table A 8: VEC Stability test

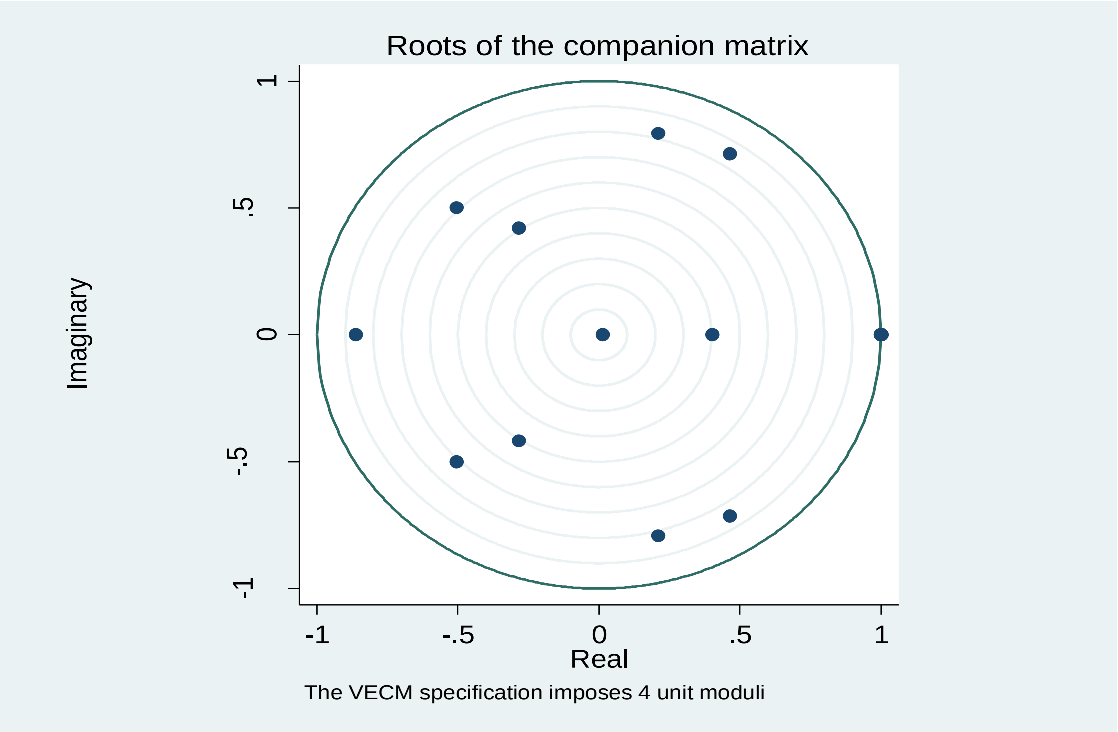
```
. vecstable, graph
```

Eigenvalue stability condition

| Eigenvalue | Modulus |
|---|---|
| 1 | 1 |
| 1 | 1 |
| 1 | 1 |
| 1 | 1 |
| -.8615289 | .861529 |
| .4647363 + .7136188*i* | .851605 |
| .4647363 - .7136188*i* | .851605 |
| .2107307 + .7928516*i* | .820379 |
| .2107307 - .7928516*i* | .820379 |
| -.5042112 + .4994107*i* | .709676 |
| -.5042112 - .4994107*i* | .709676 |
| -.2825696 + .4195614*i* | .505843 |
| -.2825696 - .4195614*i* | .505843 |
| .4023784 | .402378 |
| .01458988 | .01459 |

The VECM specification imposes 4 unit moduli.

### Figure A 1: VEC Stability test




### Table A 9: Autocorrelation test

```
. veclmar, mlag(3)
```

Lagrange-multiplier test

| lag | chi2 | df | Prob > chi2 |
|---|---|---|---|
| 1 | 13.4622 | 25 | 0.97032 |
| 2 | 31.8483 | 25 | 0.16245 |
| 3 | 16.2968 | 25 | 0.90570 |

H0: no autocorrelation at lag order

## Table A 10: Log-likelihood test for goodness of fit

```
. reg lnareac lnrprice lnyield lnrmzprice

      Source |       SS       df       MS              Number of obs =      33
-------------+------------------------------           F(  3,    29) =   23.91
       Model |  1.98703837     3  .662346123           Prob > F      =  0.0000
    Residual |   .80328555    29  .027699502           R-squared     =  0.7121
-------------+------------------------------           Adj R-squared =  0.6823
       Total |  2.79032392    32  .087197622           Root MSE      =  .16643

------------------------------------------------------------------------------
     lnareac |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    lnrprice |  -.2729431   .2019896    -1.35   0.187    -.6860583     .140172
     lnyield |  -.8748899   .1628534    -5.37   0.000    -1.207963   -.5418173
  lnrmzprice |  -.2685622   .1488373    -1.80   0.082    -.5729686    .0358443
       _cons |   18.15944   1.220444    14.88   0.000     15.66335    20.65553
------------------------------------------------------------------------------
```

## A.6: QUESTIONNAIRE

Fill the following sections with the relevant information Mumias outgrowers.

**Section I: Sugarcane details**

| Year | Area under cane by outgrowers (hectares) | Area harvested by outgrowers (hectares) | Cane delivered by outgrowers (Tons) |
|---|---|---|---|
| 1980 | | | |
| 1981 | | | |
| 1982 | | | |
| 1983 | | | |
| 1984 | | | |
| 1985 | | | |
| 1986 | | | |
| 1987 | | | |
| 1988 | | | |
| 1989 | | | |
| 1990 | | | |
| 1991 | | | |
| 1992 | | | |
| 1993 | | | |
| 1994 | | | |
| 1995 | | | |
| 1996 | | | |
| 1997 | | | |
| 1998 | | | |
| 1999 | | | |
| 2000 | | | |
| 2001 | | | |
| 2002 | | | |
| 2003 | | | |
| 2004 | | | |
| 2005 | | | |
| 2006 | | | |
| 2007 | | | |
| 2008 | | | |
| 2009 | | | |
| 2010 | | | |
| 2011 | | | |
| 2012 | | | |
| 2013 | | | |

**Section II: Rainfall and prices**

| Year | Price received per ton of cane delivered (Ksh/Ton) | Maize producer price within cane growing zone ( Ksh/90kg bag) | Consumer price index for base year 2009 (2009=100) | Mean annual Rainfall received within cane growing zone (mm) |
|---|---|---|---|---|
| 1980 | | | | |
| 1981 | | | | |
| 1982 | | | | |
| 1983 | | | | |
| 1984 | | | | |
| 1985 | | | | |
| 1986 | | | | |
| 1987 | | | | |
| 1988 | | | | |
| 1989 | | | | |
| 1990 | | | | |
| 1991 | | | | |
| 1992 | | | | |
| 1993 | | | | |
| 1994 | | | | |
| 1995 | | | | |
| 1996 | | | | |
| 1997 | | | | |
| 1998 | | | | |
| 1999 | | | | |
| 2000 | | | | |
| 2001 | | | | |
| 2002 | | | | |
| 2003 | | | | |
| 2004 | | | | |
| 2005 | | | | |
| 2006 | | | | |
| 2007 | | | | |
| 2008 | | | | |
| 2009 | | | | |
| 2010 | | | | |
| 2011 | | | | |
| 2012 | | | | |
| 2013 | | | | |