AN EXPLANATORY MODEL OF SOUTH AFRICAN YELLOW MAIZE FUTURES PRICES

S. Van Tonder*, J.H. Van Rooyen*

Abstract

This study attempts to identify the important variables that may affect yellow maize futures prices in the South African derivatives market. Data was obtained from the South African Futures Exchange, a division of the Johannesburg Securities Exchange. Weekly data on the rand-dollar exchange rates were obtained from the South African Reserve Bank (SARB). Monthly data regarding import volumes, export volumes, maize consumption and maize stocks in South Africa are available from South African Grain Information Service (SAGIS). Fifteen variables that may be used to forecast futures prices were identified from theory and similar studies. A correlation matrix of these variables with maize futures prices was determined at the 5% significance level. After applying various statistical analyses to test for autocorrelations, stationarity etc., only four variables were left with which to model the futures prices. The R2 of the remaining variables was only 12.21%, indicating a low goodness of fit. Applying the regression model to the ex-post prices clearly indicated that these variables that were identified do not adequately explain the movement in the futures prices. The primary reasons for the low accuracy of the model may be due to the use of the weather index for SA alone (a small contributor in a global market) and the linearity assumption underlying the selected dependant and independent variables may also be unrealistic. Further research is therefore needed to identify more appropriate variables with which to model yellow maize futures prices.

Keywords: Yellow Maize, Futures Prices, Regression

** Stellenbosch University, Department of Business Management, Private Bag X1, Matieland, 7602, South Africa*

Introduction and problem statement

A large degree of variability over time is present in yellow maize prices, between the periods 2006 to 2008 the yellow maize price fluctuated between R1100 and R2500 per ton (Maize Report 2010). The price variability is mainly attributable to two factors, the dependence of agricultural commodities on nature and the lag that exists between planting decisions and the harvesting of the crop (Ellis 1992).

This variability in price poses a risk to farmers, merchandisers, food processors, exporters and importers. By entering a futures contract economic agents have the ability to hedge themselves against price variability. However a successful hedge will only be achieved if the agent enters the correct position in the futures contract and the basis risk corresponding to the position in the futures contract is managed effectively (HEDGERS GUIDE). In 2009 the Chicago corn contract was launched on the Agricultural Product Division on the JSE. The contract represents the international benchmark for yellow maize prices, and since it is traded and settled in rand it is a convenient way for South African traders to hedge internationally (Commodity Derivatives JSE 2009).

Futures Markets exists primarily for hedging, but many speculators participate in the market as well. Volatility in the markets attracts speculators hoping to realize a profit by correctly anticipating and timing price movements in the futures market. These speculators play an important role to facilitate hedging since they provide liquidity to the futures market. Liquidity is defined as the ability to enter and exit the market quickly, easily and efficiently with little or no loss. This allows buyers and sellers to enter and exit a market position at an efficient price (HEDGERS GUIDE).

The use of futures goes beyond that of risk management for hedgers and trading instruments for speculators. In the South African agricultural market maize futures prices are an important determinant of the maize spot price and are therefore a price discovery mechanism. This phenomenon takes place because the majority producers and buyers prefer to fix their selling and purchasing prices respectively, in the futures market. Most of the deliveries are made in terms of futures contracts rather than in the spot market, the result is a disparity in liquidity in the spot and futures market. Therefore spot market pricing is derived from the near futures price and the fair value pricing of futures is based on the futures market itself (Faure, 2006).

Both yellow and white maize futures are traded on the South African Futures Exchange (SAFEX). The volume of maize futures traded on SAFEX has increased substantially. In 2000, 303 065 contracts were traded and this number has climbed to 1 197 216 contracts in 2008. In 2008 62.84% of all agricultural futures contracts traded on SAFEX were maize futures. The majority of maize futures traded are white maize futures which consist of 71.81% of total maize futures traded in 2008, with yellow maize consisting of 28.19% of maize futures traded in 2008. In 2008, yellow maize futures contracts were the third most traded agricultural futures contract on SAFEX, with a trading volume consisting of 17.71% of all agricultural futures traded (SAFEX, 2010).

The current and growing importance of futures in South Africa's agricultural markets, especially the maize market, is evident. Futures play an important role in risk management. Futures are very important in the sense that they are price discovery mechanisms. In addition to this, the frequent trading increases liquidity and improves the efficiency of the markets. Participants in the maize markets need to understand what factors influence the futures prices in order to effectively hedge their risk or make profitable trading decisions.

Objective of the Study

The goal of this study is to identify and quantify the factors affecting the yellow maize futures price and develop a trading model based on these factors. This model may aid traders to maximize profit and may help hedgers to set up effective risk management strategies.

Research Methodology

Multivariate statistical analysis is applied to secondary data. Using the results from the analysis a correlation matrix is generated to simultaneously identify the relationships between the independent variables. The most significant variables are identified and used in the multiple regression estimation model.

All data used in the analysis is weekly data. Corn trading data (through CBOT) concerning the independent variables have been collected from SAFEX. The South African yellow maize futures prices, trading volume and open interest on yellow maize futures at the end of every trading day were obtained since 2005. Weekly data on the rand-dollar exchange rates were obtained from the South African Reserve Bank (SARB). Monthly data regarding import volumes, export volumes, maize consumption and maize stocks in South Africa are available from South African Grain Information Service (SAGIS). Import and export parity prices have been obtained from SAGIS, the information is available on a monthly basis since 2001. Daily rainfall and temperature data in North West, Mpumalanga and Northern Cape provinces since 2005 have been sourced from the South African Weather Service.

Linear regression is used to explain the relationship between the identified independent variables and the SAFEX yellow maize near futures price. The variables included in the linear regression is identified via various statistical analyses.

Seasonality is obtained through applying the GARCH (1,1) model to futures prices for the period 1 January 2004 to 31 December 2009. This study used Microsoft Excel as a platform to apply the model to the dataset.

Daily data on temperature and rainfall in Mpumalanga, Northern Cape and North West province applicable to annual yellow maize yield is used to compile a weather index.

The focus of this study is the identification and modeling of key variables determining the market price of yellow maize futures trading on SAFEX. This study does not employ a sophisticated technique to compile a weather index. More accurate weather indices are available internationally. However, in the South African context such models do not exist. The secondary data used in this study have been obtained from various sources. There may therefore be some degree of inaccuracy and/or omissions in the data sets. Not all variables influencing yellow maize near futures price are present in the study. However, the variables considered to have the most significant impact on prices were identified and selected.

Literature review

The futures price is a forecast of what the spot price of the underlying asset of the futures contract will be for a given date in the future, based on current market information. Factors of supply and demand influencing the spot price influences the futures price in the same manner although the relationship of the changes are not always perfect (Krugel 2003). A large body of literature surrounding the factors influencing yellow maize spot and futures prices exists. In the following subsections the factors that are used in the development of the explanatory model is identified and a brief review of the existing literature surrounding these factors is provided.

Growing Conditions and a Weather indicator

Growing conditions has a direct influence on crop yields which leads to price variability. Goodwin and Schnepf (2000) found that better than average growing conditions tend to be associated with less volatile maize prices and had the strongest influence relative to other variables in their study on price variability. A study by Hennessey and Whal (1996) concluded that high temperatures together with low rainfall during growing seasons tend to increase variability in maize prices while high temperatures and high rainfall during growing seasons tend to decrease price volatility. Studies done by Chabane

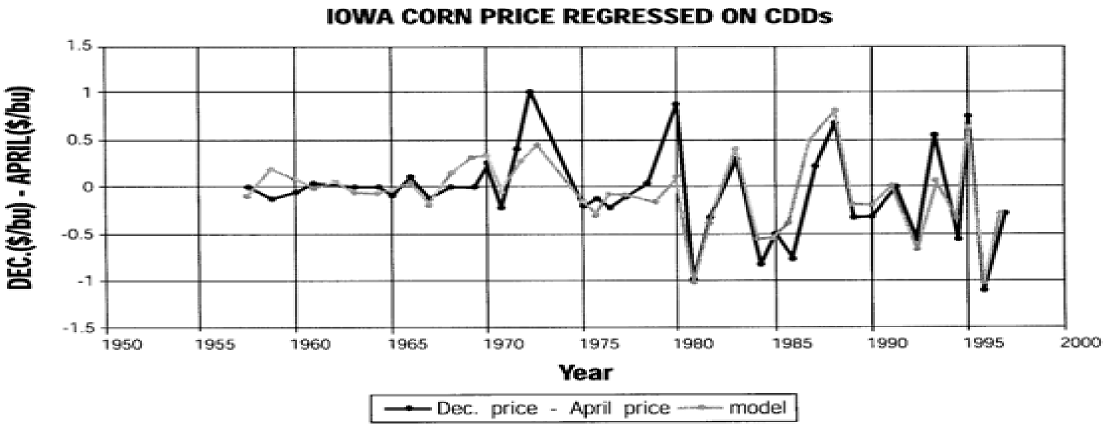
(2002) and Krugel (2003) confirms that rainfall and growing conditions have a strong influence on maize price variability in South Africa.

Auret and Schmitt (2008) suggest using the Southern Oscillation Index (SOI) as a weather indicator to predict crop yields. The SOI is a seesaw shift in air pressure at Darwin Australia and Tahiti in the South Pacific, high air pressure in Tahiti corresponds with low air pressure at Darwin (Rasmusson and Wallace 1983). El Nino conditions are indicated by sustained negative values of the SOI, this indicates a high probability of low levels of rainfall in the Southern Hemisphere. The SOI and its related indices are widely used to predict maize yields.

Heim *et al.* (2003) developed the crop Moisture Index (MSI) and the Residential Energy-Demand Temperature Index (REDTI). The MSI depicts the influence of severe drought and catastrophic wetness on maize and soybean crop yields. The REDTI provides information on the impact of seasonal temperatures on residential energy demand which is correlated to crop yields.

A study by Considine (n.d.) on the use of weather derivatives as a possible hedge for maize yields found a significant relationship between maize yields and cold degree day options (CDD). The study was conducted on Des Moines Iowa, the U.S.'s largest maize producing state between April and December for the period 1958 to 1997. Using a multiple linear regression, the April yellow maize price, the July CDD's in Des Moines, the August CDD's in Des Moines and a trend in time was modeled against the December maize price (Figure 1.1). The study concluded that a large fraction of the variability in maize prices between March and November can be explained by CCD's in Des Moines for July and August.

Figure 1. Corn Price vs. CDD's in Iowa for the Period 1958-1997



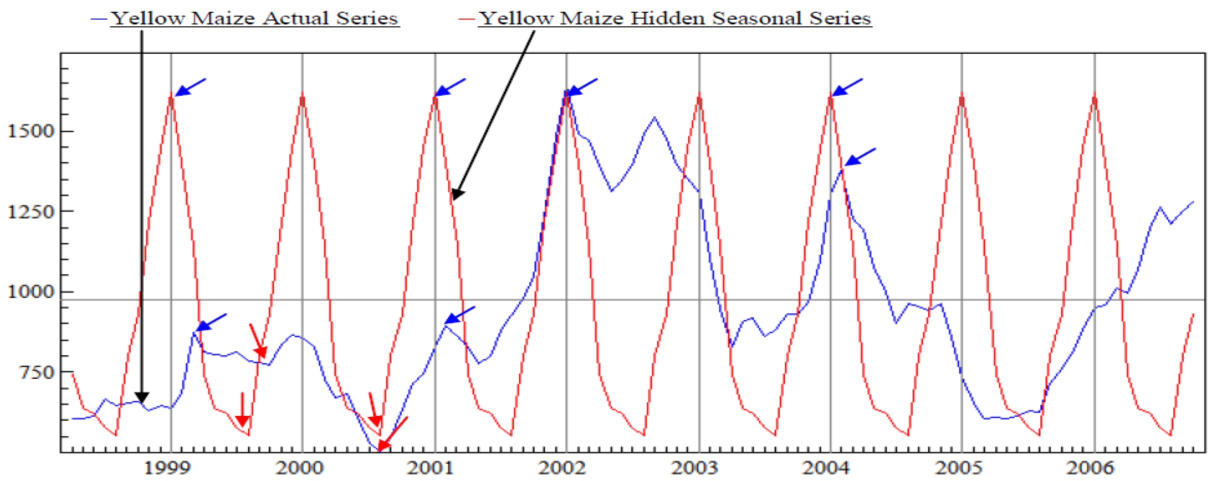
Source: Considine (n.d)

Doll (1967) suggests estimating maize yield response as a function of meteorological variables such as rainfall and temperature, trend, and their interaction. The ratio of the yield predicted for actual weather that occurred during the year to the yield had average weather occurred during the year is used to compute the index. The base yield of this ratio will change with time when interaction is present.

Seasonality

Various studies confirmed the presence of strong seasonality in grain markets arising from the seasonal nature of production technology (Fackler and Roberts 1999). Sorensen (2002) found that maize futures prices in the U.S. peak two to three months before harvest and reach their bottoms after harvesting periods. The forces of supply and demand explains this: periods of high supply are periods of relatively low prices where periods of low supply levels are also periods of relatively high prices. Goodwin and Schnepf (2000) identified strong seasonal patterns using implied volatilities calculated by the standard Black-Scholes model on maize and wheat options; however the patterns were only statistically significant in the case of maize. Their study indicates price variance peaks during summer months when maize stocks are at their lowest levels.

Figure 2. Monthly Average Spot Series of Yellow Maize Log Returns Prices versus Hidden Seasonal Series



Source: Heymans and Styger (2008)

A study conducted by Heymans and Styger (2008) suggest that seasonality should be approached as being stochastic rather than deterministic. Approaching seasonality in this manner allows for the testing of hidden components present in the time series. The study was conducted on white and yellow maize prices for the period 26 March 1998 to 13 July 2006 in the South African market using different methods to test for seasonality. They concluded that using GARCH models which treat seasonality as being deterministic provided no meaningful relationship between the results and the actual time series; however the results were statistically significant. Using the Unobserved Components Model (UCM) method, results were not statistically significant but the patterns obtained seem to describe the time series more accurately than the deterministic approach. In the case of yellow maize the unobserved seasonal patterns appears to signal the majority of changes in the yellow maize series over the test period. (Figure 2)

Methodology

Simple linear interpolation is used in cases where daily data is not available. The following equation is used:

$$Z_s = \frac{(T1-Ts)}{(T2-T1)} * Z_1 + \frac{(Ts-T1)}{(T2-T1)} * Z_2,$$

Where:
Z_s= value to be determined
T_1= date at time 1
T_s= date at which the value is to be determined
T_2= date at time 2
Z_1= given value at T_1
Z_2= given value at T_2

Microsoft Excel contains a "FORECAST" function based on the above equation that is utilized to interpolate the large datasets in the study.

Development of a Weather Index

The weather index that is used in this study is based on the relationship between temperature and rainfall and yellow maize yield. The relationship proposed by Doll 1967 is quantified as follows:

$$f_1\left[\int_0^S w(s, x_t)ds\right] = y_t$$

Weather is characterized by a function, $x_t(s)$□, representing all of the meteorological variables influencing the final crop yield in a given year, where t is the year, s is the time through the growing season, and $0 \leq s \leq S$. W is a weighting function of the meteorological variable and y_t is yield in year t.

By dividing the growing season into several periods the product becomes $w_j(s)x_{tj}(s)$, where j = 1,2,3…., k is the period and s is now the time within in a period. Yield in year t is then:

$$f_2\left[\sum_{j=1}^{k} . \int_{S_{j=1}}^{S_j} w(s)x_{tj}(s)ds\right] = y_t$$

This expression provides the foundation for estimating a weather index, the specific formulation used could vary widely. One formulation would be to assume that the integral expression for each period can be approximated by a product function such as:

$$z_{tj} = b_j x_{tj} \approx \int_{S_{j=1}}^{S_j} w_j(s)x_{tj}(s)ds$$

So that yield can be expressed as the composite function:

$$y_t = f_3\left[\sum_{j=1}^{k} z_{tj}\right] = f_3(z_t),$$

Where

$$z_t = \sum_{j=1}^{k} z_{tj}$$

b_j in this formulation corresponds to the weight function $w_j(s)$, and x_{tj} is a direct measure of some meteorological variable corresponding to the function $x_{tj}(s)$, z_{tj} provides a measure of the impact of the meteorological variable in the period j in the year t.

Assuming the z_{tj}'s are linear functions of the meteorological an empirical model can be derived from this conceptual formulation. This model would be:

$$y_t = \beta_0 + \beta_1 Z_t + \beta_2 Z^2{}_t$$

This is a quadratic function commonly used in agricultural economics research. Meteorological effects in time periods are not assumed to be independent. Weather in each period interacts with weather in every other period. An index for year t can be computed as:

$$I_t = \frac{(\beta_1 Z_t - \beta_2 Z_t^2)}{\beta_1 \overline{Z}_t + \beta_2 \overline{Z}_t^2}$$

Where $\overline{Z}_t$ and $\overline{Z}^2{}_t$ are the mean values of Z_{tj} and $Z^2{}_{tj}$ for period n. The model can be generalized to include other meteorological variables.

The meteorological variables that are used in this equation is rainfall and temperature. The above formula is programmed into Microsoft Excel and the index is calculated for the period 1 January 2004 to 31 December 2005.

Calculation of Stock-Use and Volume to Open Interest Ratios

The calculation of the stock-use ratio is defined by the following equation:

$$\frac{\textit{Beginning Stock} + \textit{Total Production} - \textit{Total Use}}{\textit{Total Use}}$$

This equation is programmed into Microsoft Excel and applied in the study in order to calculate the stock-use ratios for the period 1January 2004 to 31 December 2009.

In order to obtain the volume to open interest the following equation is utilized.

$$\frac{\textit{Total volume of open interest on futures contracts on day } t}{\textit{Total volume of futures contracts traded on day } t}$$

The equation is programmed into a Microsoft Excel spreadsheet and applied to the appropriate data in order to obtain the volume to open interest for the period 1 January 2004 to 31 December 2009.

Time to Maturity on Futures Contracts

Samuelson (1965) proposed that as a futures contract approaches maturity, futures prices will incorporate more information thus increasing price variability as time to maturity decreases; this is known as the Samuelson or maturity effect. A result from a study by Streeter and Tomek (1992) was not entirely consistent with the Samuelson effect; they found time to maturity had nonlinear effect on price variability, with price variability diminishing in the months immediately preceding maturity. This may suggest that little new information is added during the period immediately preceding contract expiration. Goodwin and Schnepf (2000) using conditional heteroscedasticity models found positive evidence supporting the Samuelson effect for wheat, however in the case of corn little evidence was found supporting the Samuelson effect. A study conducted by Smith (2005) using Partially Overlapping Time Series on simultaneously traded yellow maize futures supports the Samuelson effect. Evidence found by Duong and Kalev provides support for the Samuelson Effect in agricultural markets using the method of Seemingly Unrelated Regressions.

Export and Import Parity Prices

South Africa's maize markets operates within a free market environment, therefore changes in the world markets have direct influence on domestic prices. Calculated import and export parity prices provide price ceilings and floors respectively for domestic maize prices. A band is determined by world prices within which domestic prices can vary depending on supply and demand conditions. In the cases where demand exceeds supply, prices tend to move toward import parity prices and in cases where supply exceeds demand, prices tend to move toward export parity prices. Proagri (2001) and Krugel (2003) found that this relationship is relevant to maize futures prices in the South African context.

The Rand-Dollar Exchange Rate

The rand-dollar exchange rate is a variable in the calculation of import and export parity prices and has an influence on the price band containing maize futures prices fluctuations. Chabane (2002) found that there has been a relationship in the past between the rand-dollar exchange rate and the maize price. His study indicated that sharp currency depreciation in the rand relative to the U.S. dollar coincided with maize price increases in the past. Furthermore Vink and Kirsten (2002) established that a high elasticity exists between white maize prices and the rand dollar exchange rate, a 1% increase in the exchange rate resulted in a 1.16% increase in the real price of white maize. However this elasticity may have been exacerbated by a crop shortage during the period the study conducted. Additionally, Krugel (2003) and Auret and Schmitt (2008) found that the exchange rate has an influence on the volatility of maize prices in South Africa.

The World Price of Maize in Dollars and the CBOT near Futures Price

Auret and Schmitt (2008) suggest the world price of maize in dollar and the CBOT near futures price to be considered as a factor influencing domestic maize futures prices. It is generally accepted that the CBOT corn near futures price serves as the world supply-demand price discovery mechanism. The largest commodity exchange on which maize is traded the Chicago Board of Trade (CBOT) which is located in the U.S.. Additionally the U.S. is the largest producer, exporter and consumer of maize thus it is intuitive that changes in U.S. maize prices has an effect on the global maize market.

Analysis

Various statistical procedures and analysis is performed on the time series data concerning the variables believed to influence the yellow maize near futures prices of contracts trading on SAFEX. The following sections provides the outcomes of the statistical methods applied to the data. Before the analysis is discussed, the following assumptions regarding the analysis should be highlighted. The assumptions are as follows:

The Multiple Linear Regression Model used for the research is:

$$Y_i = \infty + \beta_j X_{ij} \dots \beta_p X_{ip} + \xi_i$$

Where Y_i ($i = 1,\dots n$) is i^{th} value of the futures price and the X_j's the j^{th} variables filtered by the above mentioned techniques.α,β_j ($j = 1,\dots p$) are the regression coefficients and reflect the partial effect of the associated variables, holding the effects of all other variables constant. p is the number of variables in the model and ξ_i is the random error term.

1. The linearity assumption – the relationships between the independent variables and the dependent variables, SAFEX yellow maize futures price, should be linear. That the regression model is linear in the coefficients β_j.
2. The assumption of normal distribution of the errors with a constant (homoscedastic) variance, meaning the error term is normally distributed with zero mean.
3. The assumptions of no autocorrelation among the errors, meaning the observations of ξi are uncorrelated with each other.
4. The assumption of no multi collinearity, meaning no independent variable is a perfect linear function of any other independent variable.

The analysis which follows is conducted under the above mentioned five assumptions. The assumptions are made to simplify the analysis of the data and may not reflect the actual relationships present among the variables in the market in a sufficiently accurate manner. The analysis commences with an analysis of the original data. A second analysis follows which is conducted on the first difference in the natural logarithm of the original data. The final model is then estimated and tested on historical data and ex-post data.

Analysis of Original Data

An exploratory analysis is conducted on the original non-transformed data. The rationale for this is to uncover inherent patterns regarding the type and strength of relationships present among the variables. A strong linear relationship between the dependant and each of the independent variables is highly desirable. A strong linear relationship among the independent variables indicates collinearity which is undesirable. Examining the correlations present among the independent variables provides a useful starting point to determine the relationships among the variables. Table 1 presents the acronyms used to identify the dependent and independent variables.

Table 1. Definition of Acronyms

Variable	Acronym
Yellow maize SAFEX near futures price	NFPYMZ
Volume to open interest	VOP
Stock use ratio	SUR
Rand-Dollar Exchange	DOLLAREX
Import Parity Prices	IPP
Export Parity Prices	EPP
White maize SAFEX near futures price	NFPWMZ
Alsi40 Index	ALSI 40
Resi20 Index	RESI 20
Weather Indicator	WEATHERIND
Seasonal Volatility	VOL
Yellow Maize free on board dollar price in Mexican Gulf	FOB US
Yellow Maize free on board dollar price in Argentinean Gulf	FOB ARG
Spread between yellow and white maize SAFEX futures prices	SPRD
Weekly yellow maize yield (per ton) in South Africa	YIELD p/t

Table 2 presents the correlation matrix between the dependant and independent variables. Significant correlation at a 5% level is indicated by "*" in the table. An examination of Table 2 indicates the SAFEX (NFPYMZ) yellow maize near futures price is significantly correlated to thirteen of the other fifteen variables. The variables are the following: VOP, IPP, EPP, SUR, SPRD, DOLLAREX, ALSI40, WEATHERIND, FOBUS, FOBARG, NFPWMZ, VOL and RESI20. Additionally table 3 gives a presentation of the correlation among the independent variables and Pearson's Spurious Correlation at a 5% significance level (indicated by *).

Table 2. Correlation Matrix Generated using Original Data

	vop	ipp	epp	sur	sprd	nfpymz	dollarex
vop	1.0000						
ipp	-0.2699*	1.0000					
epp	-0.1634*	0.9593*	1.0000				
sur	-0.4315*	0.0565	0.0297	1.0000			
sprd	0.2654*	-0.1909*	-0.0853	0.0945	1.0000		
nfpymz	-0.2290*	0.8458*	0.7796*	-0.2062*	-0.2312*	1.0000	
dollarex	-0.0738	0.6083*	0.7197*	0.0503	0.2248*	0.5263*	1.0000
alsi40	-0.1730*	0.7407*	0.6180*	-0.1810*	-0.2138*	0.8893*	0.2361*
yieldpt	-0.3756*	0.2409*	0.2252*	0.4003*	0.0304	0.0438	0.0061
weatherind	0.1544*	0.1330*	0.1470*	-0.4603*	-0.0927	0.2366*	0.1806*
fobus	-0.2190*	0.9573*	0.9068*	0.0555	-0.2213*	0.8572*	0.5267*
fobarg	-0.2292*	0.9350*	0.8775*	0.0487	-0.1857*	0.8781*	0.5268*
nfpwmz	0.0798	0.5344*	0.5583*	-0.3559*	0.0965	0.7710*	0.3890*
vol	0.2749*	-0.3274*	-0.2843*	-0.2391*	0.0495	-0.2510*	-0.3854*
resi20	-0.1619*	0.7843*	0.6695*	-0.1796*	-0.2050*	0.8580*	0.2236*

	alsi40	yieldpt	weathe~d	fobus	fobarg	nfpwmz	vol
alsi40	1.0000						
yieldpt	0.1024	1.0000					
weatherind	0.1126	-0.4639*	1.0000				
fobus	0.7776*	0.1746*	0.1332*	1.0000			
fobarg	0.8099*	0.1733*	0.1359*	0.9842*	1.0000		
nfpwmz	0.7484*	-0.0079	0.1362*	0.5588*	0.5864*	1.0000	
vol	-0.2141*	-0.1656*	0.3219*	-0.2837*	-0.2979*	0.0679	1.0000
resi20	0.9627*	0.1706*	0.1123	0.7918*	0.8094*	0.7108*	-0.2031*

	resi20
resi20	1.0000

A regression equation on the original data is estimated using the stepwise regression methodology at the 5% significance level. The purpose of the regression is to further investigate the data and establish the validity of multiple linear regressions for the variables.

The stepwise regression dropped the following variables at the 5% significance level.

- RESI20
- VOL
- FOBUS
- WEATHERIND
- YIELDp/t
- ALSI

The remaining 9 variable model has an adjusted R-squared value of 95.08% indicating a very high proportion of variability is accounted for by the model. This indicates a high goodness-of-fit for the estimated model and will be useful in assessing the overall accuracy of the model.

Table 3. Results from Linear Regression on Original Data

NFPYMZ	Coefficient	Standard error	t	p > |t|	95% Confidence Interval	
VOP	-288.2491	57.1935	-5.05	0.000	-400.9344	-175.6537
IPP	0.6024756	0.0606469	9.93	0.000	0.4830339	0.7219173
EPP	-0.8841605	0.0806249	-10.97	0.000	-1.042948	-0.7253729
SUR	-23.48178	3.097319	-7.58	0.000	-29.58183	-17.38173
SPRD	-0.6089519	0.0781208	-7.80	0.000	-0.7628077	-0.4550962
DOLLAREX	104.7496	9.729255	10.77	0.000	85.58826	123.911
FOBARG	3.713063	0.422817	8.78	0.000	2.880342	4.545785
NFPWMZ	0.4119582	0.0207611	19.84	0.000	0.37107	0.4528463
CONS	-450.865	80.35145	-5.61	0.000	-609.114	-292.6161

The Durbin Watson (DW) d-statistic measures the lack or presence of serial correlation among the errors from one observation to other observations. The ideal value for the DW statistic is 2.00 indicating the absence of autocorrelation. The DW d-statistic calculated for the 5 variable regressions is 0.4462252 indicating a strong presence of autocorrelation.

The Variance Inflation (VIF) and Tolerance (1/VIF) are used to determine if collinearity is present in the variables. A VIF > 10 and a Tolerance < 0.1 indicates that collinearity may be a problem. An examination of Table 4 indicates that collinearity is a problem for 2 of the 8 variables namely IPP and EPP.

Table 4 Variance Inflation Results on Independent Variables Using Original Data

Variable	VIF	1/VIF
ipp	33.11	0.030203
epp	24.27	0.041204
fobarg	9.82	0.101836
dollarex	2.87	0.348277
nfpwmz	2.32	0.431167
sur	1.73	0.577630
vop	1.66	0.600684
sprd	1.55	0.644825
Mean VIF	9.67	

It is important to confirm whether a series is stationary before using it in a regression. A series is defined as stationary if the mean and autocovariances of the series do not depend on time. In order to test for the stationarity of a series, unit root tests are utilized. A Dicky-Fuller Augmented (DFA) unit root test which includes a constant in the test regression is used. The automatic lag length employed by the DFA test is determined using a Schwarz Information Criterion. The Schwarz Information Criterion determined a maximum lag of 14 and the results of the DFA tests are presented in table 5.

Table 5. Augmented Dicky Fuller Unit Root Test Results on Original Data

VARIABLE	DFA	t-stats	Probability at 5% level
VOP*	-2.786	-2.88	0.063
IPP	-1.658	-2.88	0.453
EPP	-1.837	-2.88	0.3621
SUR*	-3.415	-2.88	0.014
SPRD	-2.325	-2.88	0.164
FOBARG	-1.706	-2.88	0.4278
DOLLAREX	-1.887	-2.88	0.3426
NFPWMZ	-2.724	-2.88	0.07
RESI20	-2.179	-2.88	0.2141
VOL*	-2.944	-2.88	0.0404
FOBUS	-1.752	-2.88	0.4044
WEATHERIND*	-4.299	-2.88	0.0004
ALSI40	-2.028	-2.88	0.257
YIELDp/t*	-4.538	-2.88	0.0002
NFPYMZ	-1.932	-2.88	0.3171

Transforming the Data

Undesired autocorrelation is present in the time series data. Examining Table 5 reveals that only 5 (marked with *) of the 15 original variables used in the study pass the DFA test at the 5% level namely:

- VOP
- SUR
- VOL
- WEATHERIND
- YIELDp/t

A method to address the issue of autocorrelation in the original data is to transform the data by taking the first difference in the natural logarithm of the data. The correlation matrix presented in Table 6 presents the correlations among the transformed values. Significant correlations are indicated by * in the table.

Table 6. Correlation Matrix Generated Using Transformed Data

	vop	ipp	epp	sur	sprd	nfpymz	dollarex
vop	1.0000						
ipp	-0.0489	1.0000					
epp	-0.0047	0.6563*	1.0000				
sur	0.1864*	0.0190	-0.0132	1.0000			
sprd	-0.0504	0.0574	0.0588	-0.0456	1.0000		
nfpymz	-0.0503	0.2716*	0.2325*	-0.0720	0.0933	1.0000	
dollarex	-0.0354	0.2043*	0.1471*	-0.0244	0.1551*	0.0108	1.0000
alsi40	-0.0182	0.0490	-0.0447	-0.0693	-0.1551*	0.0124	-0.1310*
yieldpt	-0.1151	0.1465*	0.0935	0.0968	-0.0473	0.0307	-0.1122
weatherind	0.0405	0.0200	0.0511	0.0246	0.0182	0.0742	-0.1097
fobus	0.1076	0.1527*	0.0912	0.0549	-0.0415	0.1527*	-0.1083
fobarg	0.0484	0.1234*	0.0726	0.0368	-0.0224	0.1937*	-0.1259*
nfpwmz	0.0415	0.0913	0.0261	-0.0582	-0.0192	0.1050	0.0214
vol	0.2936*	-0.0282	-0.0060	0.6005*	-0.0100	0.0764	-0.1423*
resi20	-0.0604	0.0912	0.0932	0.0070	-0.1052	0.0320	-0.1401*

	alsi40	yieldpt	weathe~d	fobus	fobarg	nfpwmz	vol
alsi40	1.0000						
yieldpt	0.0900	1.0000					
weatherind	0.0034	-0.0579	1.0000				
fobus	0.0513	0.1018	0.0332	1.0000			
fobarg	0.0871	0.1152	0.0744	0.8684*	1.0000		
nfpwmz	0.0877	0.0059	0.0312	0.0039	0.0105	1.0000	
vol	0.0012	0.0223	0.0570	-0.0217	-0.0153	0.0354	1.0000
resi20	0.0532	0.0854	0.0135	0.0425	0.0786	-0.0130	0.0003

	resi20
resi20	1.0000

The dependant variable namely NFPYMZ is correlated to the following independent variables at a 5% significance level: IPP, EPP, FOBUS, and FOBARG.

In contrast to the large collinearity present in the original data, the transformed data displays low levels of collinearity. For further investigation, the stepwise regression method at the 5% significance level is conducted on the transformed data. The regression eliminated the following variables with p >= 0.05:

- VOP, p = 0.219
- EPP, p = 0.281
- SPRD, p = 0.273
- DOLLAREX, p = 0.927
- ALSI40, p = 0.748
- YIELDp/t, p = 0.797
- WEATHERIND, p = 0.489
- FOBUS, p = 0.669
- FOBARG, p = 0.067
- NFPWMZ, p = 0.249
- RESI20, p = 0.962

A linear regression is conducted using the remaining 3 variables namely VOL, SUR and IPP. The 3 variable regression has an Adjusted R-squared of 9,59% indicating a low degree of variability is accounted for by the model. The DW d-statistic is 1.674887 indicating the presence of moderate autocorrelation. In comparison to the 8 variable regression on the original data, the 3 variable regression has a lower goodness-of-fit of 9.59% as opposed to 95.08% mentioned before. However, the DW d-statistic has improved significantly from 0.4462252 to 1.674887. Additionally, the mean VIF for the 3 variable regression is 1.38 which is a significant improvement relative to the 9.67 of the 8 variable regression. In conclusion, using the transformed data addressed the issue of collinearity but the Adjusted R-squared (goodness-of-fit) concurrently declined.

Lastly the DFA unit root test for stationarity is conducted on the transformed data. The DFA test has the same parameters as the DFA test conducted on the original data. As indicated in Table 8, all 15 transformed variables pass the DFA unit root test at the 5% significance level. This indicates a vast decline in autocorrelation relative to the 10 untransformed variables which did not pass the DFA unit root test at the 5% significance level.

Table 7. Augmented Dicky Fuller Unit Root Test Results on Transformed Data

VARIABLE	**DFA**	**t-stats**	**Probability at 5% level**
VOP	-5.625	-2.88	0
IPP	-2.882	-2.88	0.0475
EPP	-3.13	-2.88	0.0244
SUR	-4.535	-2.88	0.0002
SPRD	-3.024	-2.88	0.0327
NFPYMZ	-3.414	-2.88	0.0105
DOLLAREX	-3.948	-2.88	0.0017
ALSI40	-3.366	-2.88	0.0122
YIELDp/T	-4.965	-2.88	0
WEATHERIND	-4.187	-2.88	0.0007
FOBUS	-3.239	-2.88	0.0178
FOBARG	-3.611	-2.88	0.0055
NFPWMZ	-3.094	-2.88	0.027
VOL	-4.292	-2.88	0.0005
RESI20	-3.275	-2.88	0.0161

The Final Model

It was decided to include FOBARG as a fourth variable in the model. The stepwise regression dropped FOBARG at a 5% significance level (0.067 > 0.05) but relative to the other dropped variables the p-value for FOBARG is low, motivating the variables' inclusion in the model. The new 4 variable linear regression displays the following results.

Table 8. Linear Regression on Final Transformed Data

```
. regress nfpymz ipp sur vol fobarg
```

Source	SS	df	MS
Model	.040017463	4	.010004366
Residual	.25489771	254	.001003534
Total	.294915173	258	.001143082

Number of obs =	259
F(4, 254) =	9.97
Prob > F =	0.0000
R-squared =	0.1357
Adj R-squared =	0.1221
Root MSE =	.03168

nfpymz	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ipp	.4383707	.0990744	4.42	0.000	.2432589	.6334826
sur	-.0229909	.0079514	-2.89	0.004	-.0386501	-.0073318
vol	.028248	.00968	2.92	0.004	.0091848	.0473112
fobarg	.134053	.0457166	2.93	0.004	.0440212	.2240848
_cons	.0016282	.0019906	0.82	0.414	-.0022919	.0055483

Studying Table 8 reveals the 4 variable regression has a higher Adjusted R-squared value of 12.21% relative to the 9.59% Adjusted R-squared value of the 3 variable model, indicating an improvement in the goodness-of-fit of the model. Additionally the DW d-statistic for the 4 variable regression is 1.71 which implies lower autocorrelation is present relative to the 3 variable regression.

Including FOBARG in the regression yields a lower mean VIF relative to the 3 variable regression, reinforcing the reduction in collinearity due to its inclusion in the regression. To summarize, the final model has an Adjusted R-squared value of 12.21% indicating the model does not account for a material proportion of variability. The DW d-statistic is 1.71 which is close to 2 indicating autocorrelation is present but not a concern. Collinearity is not present among the independent variables, since the VIF for all independent variables are less than 10.

Based on the parameter estimates for yellow maize, the following model was constructed:

First difference in the log of SAFEX Yellow Maize Near Futures Price (NFPYMZ):

= 0.0016282 + 0.4383707(1stdlogIPP) - 0.0229909(1stdlogSUR) + 0.028248(1stdlogVOL) + 0.134053(1stdlogFOBARG)

A multiple regression equation was estimated using transformed variables to account for the autocorrelation effect present in the original data, meaning the kth variable at time t was transformed into:

$X^*_{tk} = (\log X_{tk} - \log X_{tk-1})$ and the dependant variable also into $Y^*_k = (\log Y_t - \log Y_{t-1})$

Thus the fitted regression equations are in the form:

$$Y^*_i = \alpha + \sum_j \beta_j X^*_{ij} + \xi_i$$

This model is evaluated (using the transformed independent variables) and Y^*_t estimated then the predicted value Y_t at the time t would recessively be given by:

$$Y_t = (\log Y_t - \log Y_{t-1})$$

Figure 3. Model Futures Prices vs. Actual Historical Futures Prices

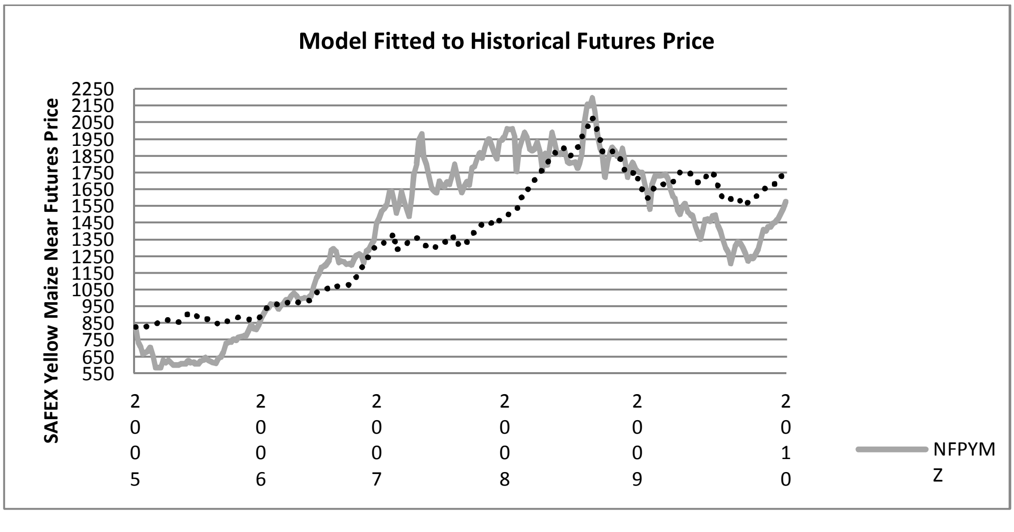
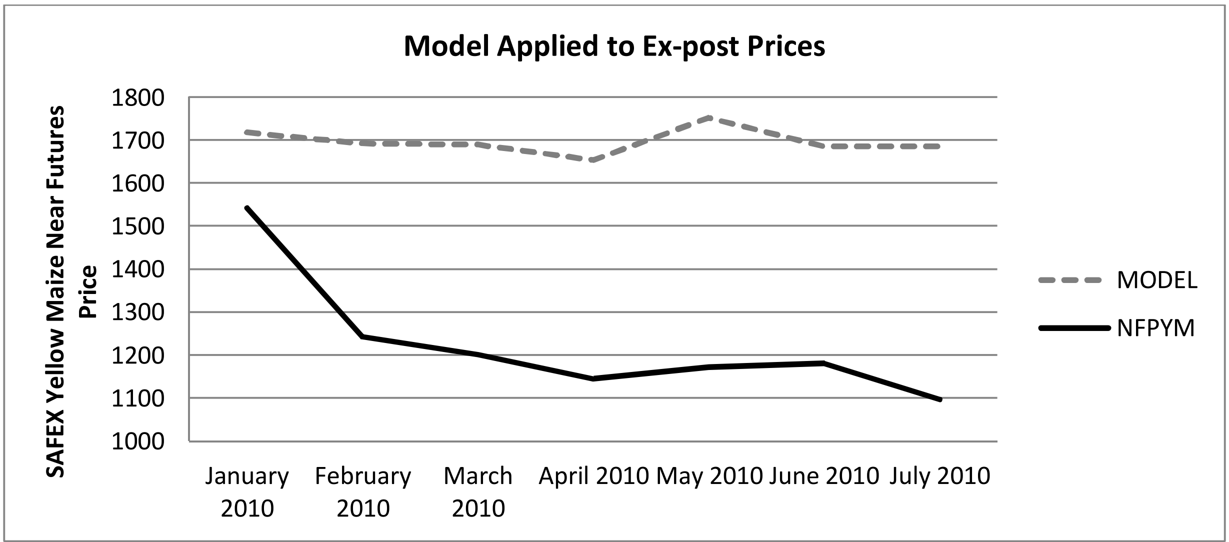


Figure 3 illustrates the actual yellow maize near futures price against the results generated by the final model for the period 1 January 2005 to 31 December 2009, a low goodness of fit is clearly visible, this is to be expected due to the Adjusted R-Squared value of 12.21% present in the regression.

Ex-post Results

The model was tested ex-post for 6 months after December 2009. Figure 4 displays the ex-post results for the 6 month period with the usefulness of results being debatable. The model gives a poor indication of the magnitude and direction of price changes.

Figure 4. Final Model Fitted to Ex-post Prices



Conclusions and Suggestions for Future Research

The independent variables used in this study do not provide an adequate explanation when modeled to determine the SAFEX yellow maize futures price. There are three possible explanations for this inadequacy. Firstly the autocorrelation among the selected independent variables are present at undesirably high levels. Secondly the proportion of variables relevant only to South Africa is too high. The weather index which only captures growing conditions within South Africa, South African yellow maize yield and the spread between SAFEX yellow and white maize futures are examples of independent variables unique to South Africa. The market for yellow maize is a vast global market and South Africa's contribution and influence on the market is minute relative to the larger participants such as Argentina and the U.S.. Thirdly the assumption of linear relationships between dependent and independent variables is unrealistic and may also contribute to the unrealistic outcome.

Further research may be conducted especially as it relates to the selection of variables for inclusion in the model. In order to more accurately quantify the effect of growing conditions on maize yields and prices, growing conditions should be analyzed in areas producing a significant volume of global yellow maize to the market. The relaxation of the linearity assumption through the use of non-linear regression may improve the accuracy of results. Finally, market momentum induced by trading sentiment may influence futures prices in a stochastic manner which cannot be accounted for by any deterministic variable. Future estimation models should take this into account.

References

1. Auret. C.J., Schmitt. C.C. 2008. An Explanatory Model of South African White Maize Futures Prices. Journal for Studies in Economics and Econometrics, 32(1): 103-131.
2. Chabane, N. 2002 An Evaluation of the Influences on Price and Production in the Maize Market following Liberalisation. A paper for the Trade and Industrial Policy Secretariat Annual Forum 19-22 September, Glenburn Lodge, Muldersdrift, South Africa. [Online]. Available: http://www.tips.org.za/files/566.pdf. [10/10/2010].
3. Considine, G. n.d. Introduction to Weather Derivatives. Weather Derivatives Group,
4. Duong, H. N. & Kalev, P.S. 2008. A test of the Samuelson Hypothesis using Realized Range. Journal of Futures markets, 28(7):680-696. [Online] Available: http://www3.interscience.wiley.com/journal/119030215/abstract?CRETRY=1&SRETRY=0. [09/02/2010].
5. Doll, J.P. 1967. An Analytical Technique for Estimating Weather Indexes from Meteorological Measurements. Journal of Farm Economics,49(1):79-88. [Online] Available: http://www.jstor.org/stable/1237069. [17/02/2010].

6. Fackler, P.L & Roberts, M.C. 1999. A Term Structure Model for Agricultural Futures. Selected Paper prepared for the American Agricultural Economics Association Meetings. [Online]. Available: http://ageconsearch.umn.edu/bitstream/21543/1/sp99fa02.pdf. [10/02/2010].
7. Faure, A.P. 2006. Fair Value Pricing of Agricultural Futures in South Africa. South African Journal of Economics, 74(2):261-265. [Online]. Available: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1813-6982.2006.00069.x. [08/02/2010].
8. Goodwin, BK & Schnepf, R. 2000. Journal of Futures Markets. 20(8): 753–774, September 2000.
9. Hennessey, D. A., & Wahl, T. I. 1996. The effects of decision making on futures price variability. American Journal of Agricultural Economics, 78(3):591–603. [Online]. Available: http://www.jstor.org/stable/1243277. [09/02/2010].
10. Heim, R.R., Lawrimore, J.A., Wuertz, D.B., Waple, A.M., Wallis, T.W.R. 2003: The REDTI and MSI: Two New National Climate Impact Indices. 42: 1435–1442.
11. Heymans, A. & Styger, P. 2008. Seasonality As An Unobservable Component In South African Agricultural Market Data. International Business & Economics Research Journal, 397(7). [Online] Available: www.cluteinstitute-onlinejournals.com/PDFs/638.pdf. [10/02/2010].
12. Krugel, L.J. 2003. Some Factors Influencing the Price of White Maize Futures Contracts in South Africa. Paper prepared for the Biannual Conference of the Economic Society of South Africa. Somerset West, South Africa. [Online]. Available: http://www.ekon-oom.com/hot%20topics/Krugel_Maize%20futures.pdf. [08/02/2010].
13. South African Futures Exchange. 2010. [Online]. Available: *www.safex.co.za.* [17/02/2010].
14. Proagri. 2001. Maize Industry Outlook. [Online] http://www.proagri.co.za/uitgawe_35/35-14_FNB.htm. [08/02/2010].
15. Rasmusson, E.M. & Wallace, J.M. 1983. Meteorological Aspects of the El Nino/Southern Oscillation. Science. 222(4629):1195-1202 [Online] Available: http://www.sciencemag.org/cgi/content/abstract/222/4629/1195. [12/02/2010].
16. Sorensen, C. 2002. Journal of Futures Markets. 22(5): 393–426, May 2002.
17. Streeter, D. H., & Tomek,W. G. 1992. Variability in soybean futures prices: An integrated framework. Journal of Futures Markets. 12(6):705–728. [Online]. Available: http://www3.interscience.wiley.com/journal/112772453/abstract. [9/02/2010].
18. Vink, N. & Kirsten, J. 2002. Pricing Behaviour in the South African Food and Agricultural Sector. A report to the National Treasury of South Africa. [Online]. Available:http://www.sarpn.org.za/documents/d0000327/P280_Pricing.pdf. [13/02/2010].