Model Development for South African Producer Price of Maize (1970-2001)

John Rutter
University of Tennessee
Statistics 571

July 31, 2003

Model Development for South African Producer Price of Maize

Introduction

Within the world market for maize (corn), the United States is the dominant exporter. Argentina and South Africa are also large exporters of maize but have a smaller market share. The price of maize can be driven by internal or external factors depending on whether the market is open-unregulated or closed-regulated, as well as the amount of market share. Until recently, the South African producer price for maize was regulated. Therefore, the hypothesis is that the internal factors for a country were more influential in the price determination for the regulated period than external factors, which are used in the determination of the elasticity of price transmission. The hypothesis will be tested using multiple regressions with the South African producer price of maize as the response variable. Internal and external factors will be analyzed as possible predictors for the response variable. The study will analyze South African producer prices from 1970 to 2001.

Maize is important to producers as well as the people of South Africa. Yellow and white maize account for 35% of the total cultivated area. Figure 1 displays a map with the area of maize production, which is located in the central and eastern part of the country. Furthermore, white maize is the staple food crop for the country. The South African maize market was regulated by the Maize Board from 1968 to 1997 in accordance with the Agricultural Marketing Act of 1968. The Maize Board would set the price of maize in April of each year before the harvesting and marketing of the crop. The price was developed using the previous year's price and supply/demand conditions[1].

Based on literature on the South African Maize Market, the Maize Board provided a "one-channel market," which is highly regulated[2]. Therefore one would suspect that internal factors, such as price as well as supply and demand, would be more important than external factors, such as world price. For this reason, the scope of this model is expanded to determine what predictors are important to the response variable or the South African producer price of maize.

[1]Essinger, Stacy M., Lowell D. Hill, and Jacobus M. Laubscher. "Privatization Progress in the South African Maize Industry." June 1998. University of Illinois at Urbana-Champaign. AE-4723. http://www.ace.uiuc.edu/research/ae/papers/ae4723.pdf

[2] ibid.

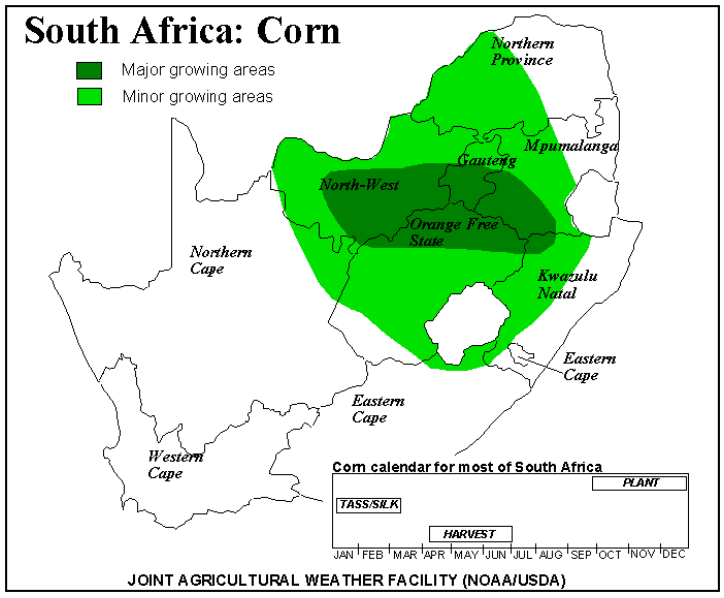


Figure 1. Area of Production of South African Maize
Source: Spectrum Commodities

Data

The producer prices for South African maize and the area planted in maize were obtained from Walter Moldenhauer, Directorate: Agricultural Statistics. Figure 2 provides a scatter plot of the data. An initial observation of the data shows that a transformation is needed to linearize the data. Figure 3 provides a scatter plot of the data with a natural logarithm transformation.

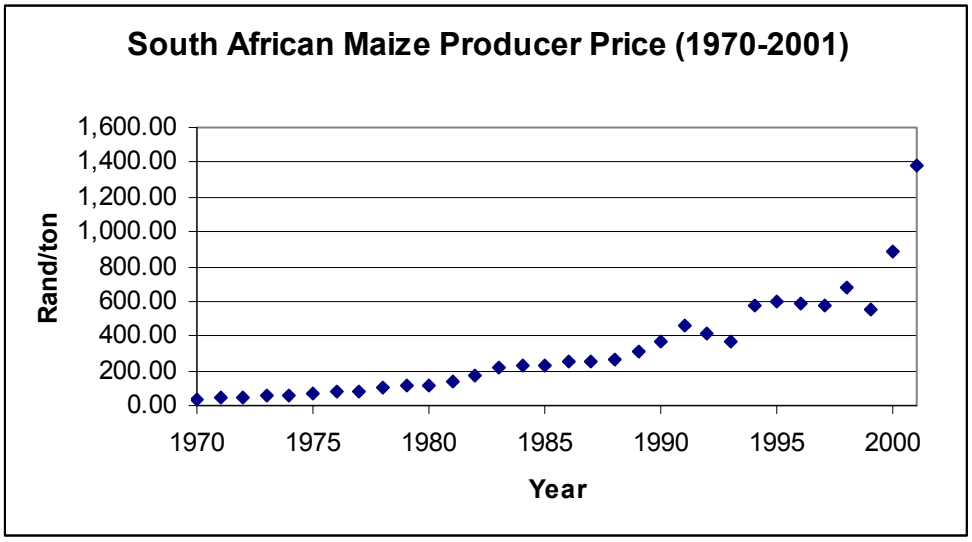


Figure 2. Scatter Plot

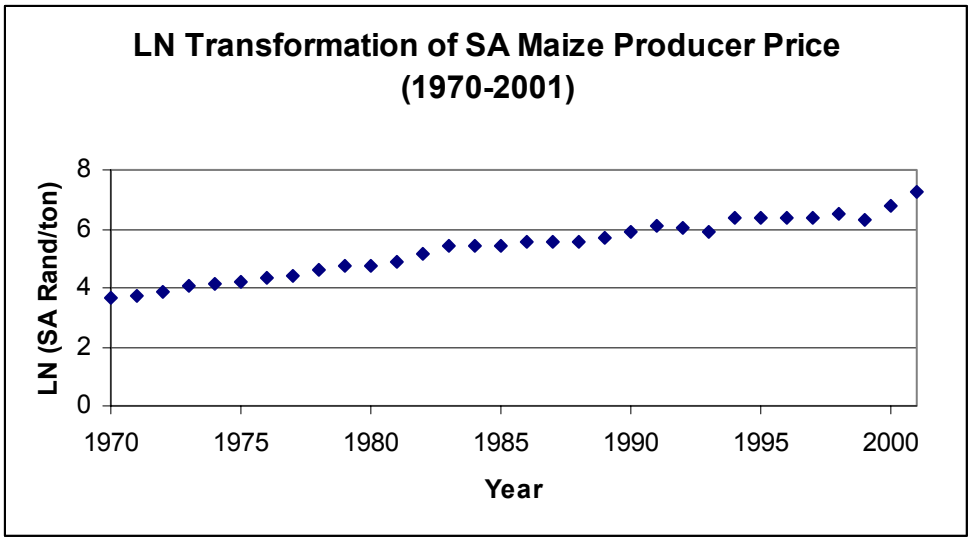


Figure 3. Scatter Plot with Natural Logarithm Transformation

In addition to the response variable, South African producer price of maize, and area planted, the other predictor variables were obtained through the United States Department of Agriculture and more directly the Economic Research Service.

Methodology

The objective of this study is to evaluate the appropriate predictor values for the response variable. Table 1 shows the variables considered in the model as well as the units. The response variable is natural logarithm of p_{sa}. Preliminary analysis showed two distinct periods, which include the operation of the Maize Board and the deregulated market (without Maize Board). The effect of this policy change provided a period of the Maize Board from 1970-1996 and a period without the Maize Board from 1997 to present. The variables listed in Table 1 were placed in a Stepwise Regression for the period with the Maize Board and picked based on the significance of the variable and the percent variation explained in the model. The lagged South African Producer Price of Maize was also transformed using the natural logarithm to linearize the data.

Table 1 provides a description of each of the variables considered in the model. Further explanation of the dummy variables may be deemed appropriate. Early analysis of the model indicated that 1972 was a possible outlier. In 1972, the U.S. sold Russia 440 million bushels of wheat. The impact of this sale raised not only wheat prices but also feed grain prices[3]. This is suspected to be the cause of the 1972 outlier, because South Africa was a feed grain exporter at this time. In addition, a dummy variable was also applied to the Maize Board and evaluated. Since the dummy variable would effectively eliminate the final 5 years, it was simpler to divide the data into two periods. Problems also resulted when the dummy variable was applied to the regression. The dummy variable that affected the intercept was not significant, and a dummy

[3] Luttrell, Clifton B. "The Russian Wheat Deal – Hindsight vs. Foresight." *Fiederal Reserve Bank of St. Louis*. October 1973. pp. 2-9. http://research.stlouisfed.org/publications/review/73/10/Russian _Oct1973.pdf

variable that affected the slope and intercept through an interaction variable was significant but contained multicollinearity. Therefore, the dummy variable was applied to the Russian deal and the data was separated into two periods, which this paper focuses on the period with the Maize Board.

An understanding of the yield variable is also important to the analysis. Yield can be affected by the genetics in the seed and also by the weather. Therefore, weather is considered to be a confounding variable in the analysis. This impact is recognized in the model.

Abbreviation	**Variable**	**Units**
SA or p_{sa}	South African Maize Producer Price	Rand / ton
SA-1 or p_{sa-1}	Lagged South African Maize Producer Price	Rand / ton
US	U.S. Maize Producer Price	Rand / ton
Year	Year	1970-2001
Rand/$US	South African to U.S. Exchange Rate	Rand / $US
$US/Rand	U.S. to South African Exchange Rate	$US / Rand
Area Planted	Area Planted in South Africa	1000 ha
Yield	Yield in South Africa	mt / ha
Exports	South African Exports	1000 mt
STU	South African Stock-to-Use Ratio	ratio
Maize Board	Maize Board Operation from 1968 to 1997 (Dummy)	0 or 1
Russian Deal	Russian Deal (Dummy)	0 or 1

Table 1. Model Variables Considered in the Study

Figure 4 displays the results of the Stepwise Regression. In the Step History, the lagged natural log of the South African price explains almost 98% of the variation in the response variable. This is followed by yield and area planted, which impact the supply of maize. The Russian Deal dummy variable is also important in explaining the impact of the Russian Deal with the U.S. For the model as a whole, demand variables like exports were not significant and excluded from the model. This dummy variable accounts for the change in demand that is present in this year. Each of the variables meets the significance criteria of 0.001.

Stepwise Fit

Response: Log SA

Stepwise Regression Control

Prob to Enter 0.250
Prob to Leave 0.100
Direction Forwa
Rules: Comb

6 rows not used due to missing values.

Current Estimates

SSE	DFE	MSE	RSquare	RSquare Adj	Cp	AIC
0.0942444	21	0.0044878	0.9945	0.9934	5.0004129	-136.119

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
☑	☑	Intercept	2.06086189	1	0	0.000	1.0000
☐	☑	SA-1 Log Rand/ton	0.89374997	1	7.506937	1672.733	0.0000
☐	☐	US Rand/ton	0	1	0.001627	0.351	0.5600
☐	☐	Year	0	1	0.012423	3.037	0.0968
☐	☐	Rand/$US	0	1	0.001072	0.230	0.6367
☐	☐	$US/Rand	0	1	0.00047	0.100	0.7547
☐	☑	Area Planted	-0.000254	1	0.142901	31.842	0.0000
☐	☑	Yield	-0.1321235	1	0.153834	34.278	0.0000
☐	☐	Exports	0	1	0.00191	0.414	0.5274
☐	☐	STU	0	1	0.007104	1.630	0.2163
☐	☐	Maize Board	0	0	0	.	.
☐	☐	Maize Board*(SA-1 Log Rand/ton-5.07622)	0	0	0	.	.
☐	☐	(Yield-1.99731)*(Area Planted-4480.23)	0	1	0.002637	0.576	0.4568
☐	☑	Russian Deal	-0.3833065	1	0.08462	18.855	0.0003

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p
1	SA-1 Log Rand/ton	Entered	0.0000	16.75743	0.9799	54.671	2
2	Yield	Entered	0.0068	0.095655	0.9855	35.356	3
3	Area Planted	Entered	0.0078	0.069559	0.9895	21.856	4
4	Russian Deal	Entered	0.0003	0.08462	0.9945	5.0004	5

Figure 4. Stepwise Regression

Results

The Stepwise Regression led to the development of the following model for the South African Producer Price for Maize in Equation 1. The natural logarithm transformation is denoted by ln. Figure 5 displays a plot of the actual versus the predicted ln p_{sa} . Based on the plot, the predictor variables do a good job of explaining the response variable.

$$\ln p_{sa} = 2.0608619 + 0.89375 * \ln p_{sa-1} - 0.132124 * yield - 0.000254 * area - 0.383306 * Russian + e_t$$

Equation 1

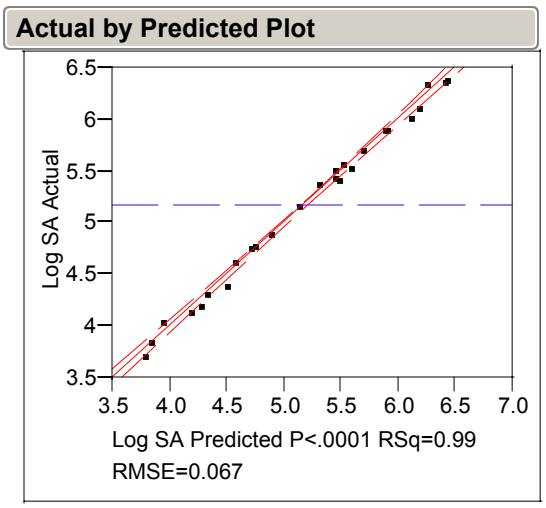
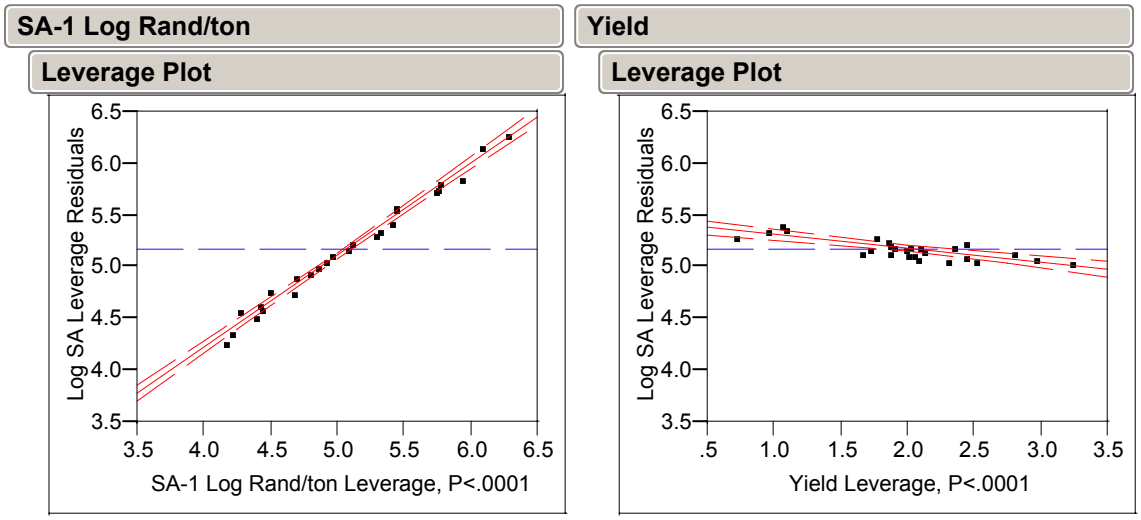


Figure 5. Actual vs. Predicted Plot

Figure 6 displays the leverage plots for each of the predictor variables. Leverage plots show from a graphical perspective the significance of the variable. Possible outliers could also be identified from this view. The leverage plots for this data do not display any distinct outliers and display the significance of each predictor in the model. The confidence intervals are also very tight for the variables, especially the lagged natural log of South African producer price of maize.



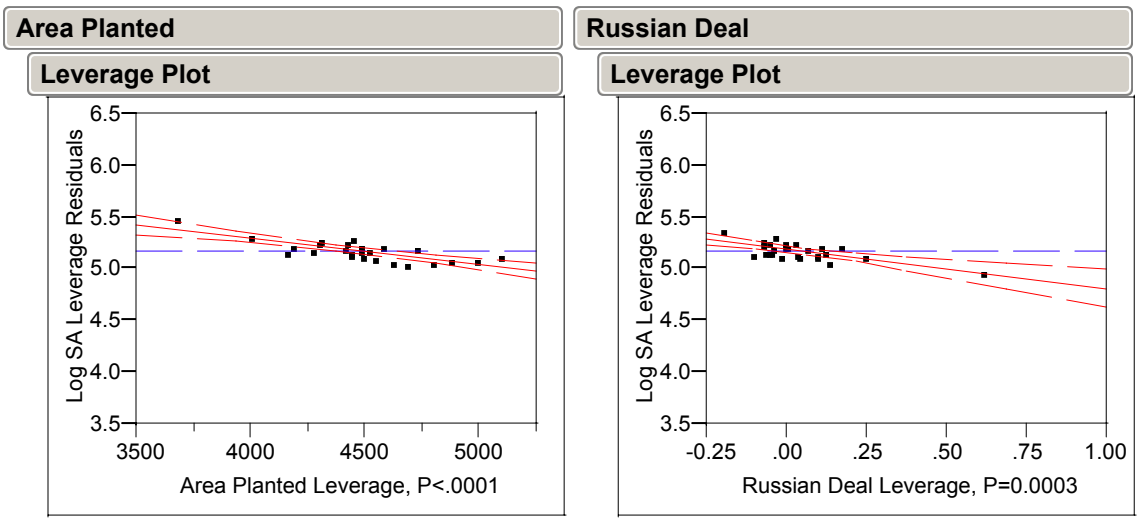


Figure 6. Leverage Plots of Predictor Variables

The proportion of the variability in the natural logarithm of the South African producer price that is accounted for by the lagged natural logarithm of the South African producer price, yield, area planted and the Russian Deal is 0.994489 or 99.4489%. The estimated variance, s^2, is 0.00449 and the estimated standard deviation, s, is 0.066991.

The goodness of fit can be analyzed using the F Test, which is as follows:

$$H_0 : \beta_1 = \beta_2 = \beta_k = 0 \text{ vs. } H_1 : \text{at least one } \beta_j \neq 0.$$

Reject the null hypothesis if $F = 947.4103 > f_{4,21,0.01} = 4.37$.

Based on the F test, reject the null hypothesis and conclude that the parameter estimate is not equal to zero for at least one predictor.

Summary of Fit

RSquare	0.994489
RSquare Adj	0.993439
Root Mean Square Error	0.066991
Mean of Response	5.181211
Observations (or Sum Wgts)	26

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	17.007260	4.25182	947.4103
Error	21	0.094244	0.00449	Prob > F
C. Total	25	17.101505		<.0001

Figure 7. Summary of Fit and Analysis of Variance

Figure 8 contains the parameter estimates for the model. Hypothesis test were developed to test each predictor. The significance level used in this analysis is 0.001. The test is as follows:

$H_0 : \beta_i = 0$ vs. $H_1 : \beta_i \neq 0$ where i is the intercept and predictor variables.
Reject H_0 at $\alpha = 0.001$ level if t Ratio > $t_{24,0.0005} = 3.745$.

For the intercept and each of the predictor variables, the null hypothesis is rejected, denoting that parameter estimates are not equal to zero. The prob > $|t|$ also shows that each of the parameter estimates is significant at the 0.001 level.

Figure 8 also displays confidence intervals and the variance inflation factor. Based on the upper and lower 95% confidence intervals, none of the confidence intervals include zero, which would indicate that none of the predictors should be excluded from the model. The model does not contain multicollinearity either. This is indicated by variance inflation factors less than 5.

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%	VIF
Intercept	2.0608619	0.299121	6.89	<.0001	1.4388055	2.6829183	.
SA-1 Log Rand/ton	0.89375	0.021853	40.90	<.0001	0.848305	0.9391949	1.8836202
Yield	-0.132124	0.022567	-5.85	<.0001	-0.179054	-0.085193	1.1226417
Area Planted	-0.000254	0.000045	-5.64	<.0001	-0.000348	-0.00016	1.8453439
Russian Deal	-0.383306	0.088273	-4.34	0.0003	-0.566881	-0.199732	1.669504

Figure 8. Parameter Estimates

The Type I and Type III Tests, as shown in Figure 9, demonstrate that each predictor are significant at the $\alpha = 0.001$. Therefore, each predictor provides additional information about the natural logarithm of South African producer price of maize that is useful in the model.

Sequential (Type 1) Tests

Source	Nparm	DF	Seq SS	F Ratio	Prob > F
SA-1 Log Rand/ton	1	1	16.757426	3733.972	<.0001
Yield	1	1	0.095655	21.3144	0.0001
Area Planted	1	1	0.069559	15.4995	0.0008
Russian Deal	1	1	0.084620	18.8554	0.0003

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
SA-1 Log Rand/ton	1	1	7.5069375	1672.733	<.0001
Yield	1	1	0.1538339	34.2780	<.0001
Area Planted	1	1	0.1429010	31.8419	<.0001
Russian Deal	1	1	0.0846197	18.8554	0.0003

Figure 9. Type I and III Sum of Squares Test

In the preliminary data analysis stages, it was determined that the South African producer price of maize needed a natural logarithm transformation. The Box Cox Transformation in Figure 10 demonstrates that this is, in fact, the proper transformation, since lambda is approximately equal to one, indicating that the data is already in the proper form.

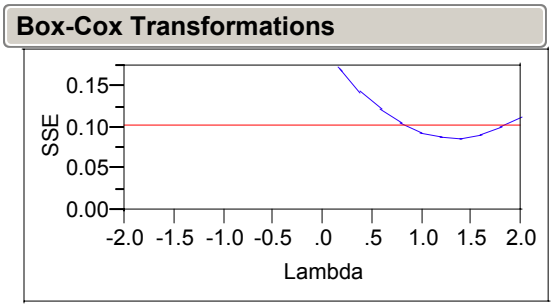


Figure 10. Box Cox Transformation

In order to evaluate if autocorrelation is present in the model, the Durbin-Watson Statistic is observed. The autocorrelation coefficient is -0.0332, which demonstrates that the model has only a little negative autocorrelation. The accuracy of this test with the use of a lagged dependent variable could be questionable.

Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
2.0165806	26	-0.0332	0.3985

Figure 11. Durbin-Watson Statistic

After the model is fit, it is important to evaluate the randomness and the normality of the residuals. Randomness in the residuals indicates that the predictors explained most of the pattern in the model. Figure 12 and Figure 13 shows that the residuals are very random for the predicted values and for the predictors. The second criteria are normality in the residuals. It is desired that the residuals be normally distributed. Figure 14 shows that the residuals are normally distributed. The Normal Quantile Plot has a linear pattern. In addition, the Shapiro-Wilk Test can be used to evaluate the normality with the null hypothesis that the residuals are normal and the alternative hypothesis that the residuals are not normal. The attained significance level is 0.4232, which is by no means small, so we do not reject the null hypothesis and conclude that the data are normal.

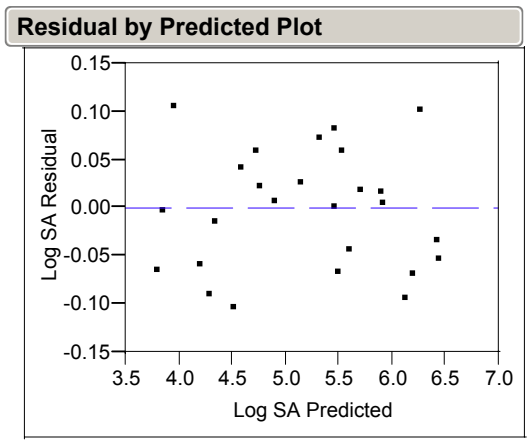


Figure 12. Residual vs. Predicted

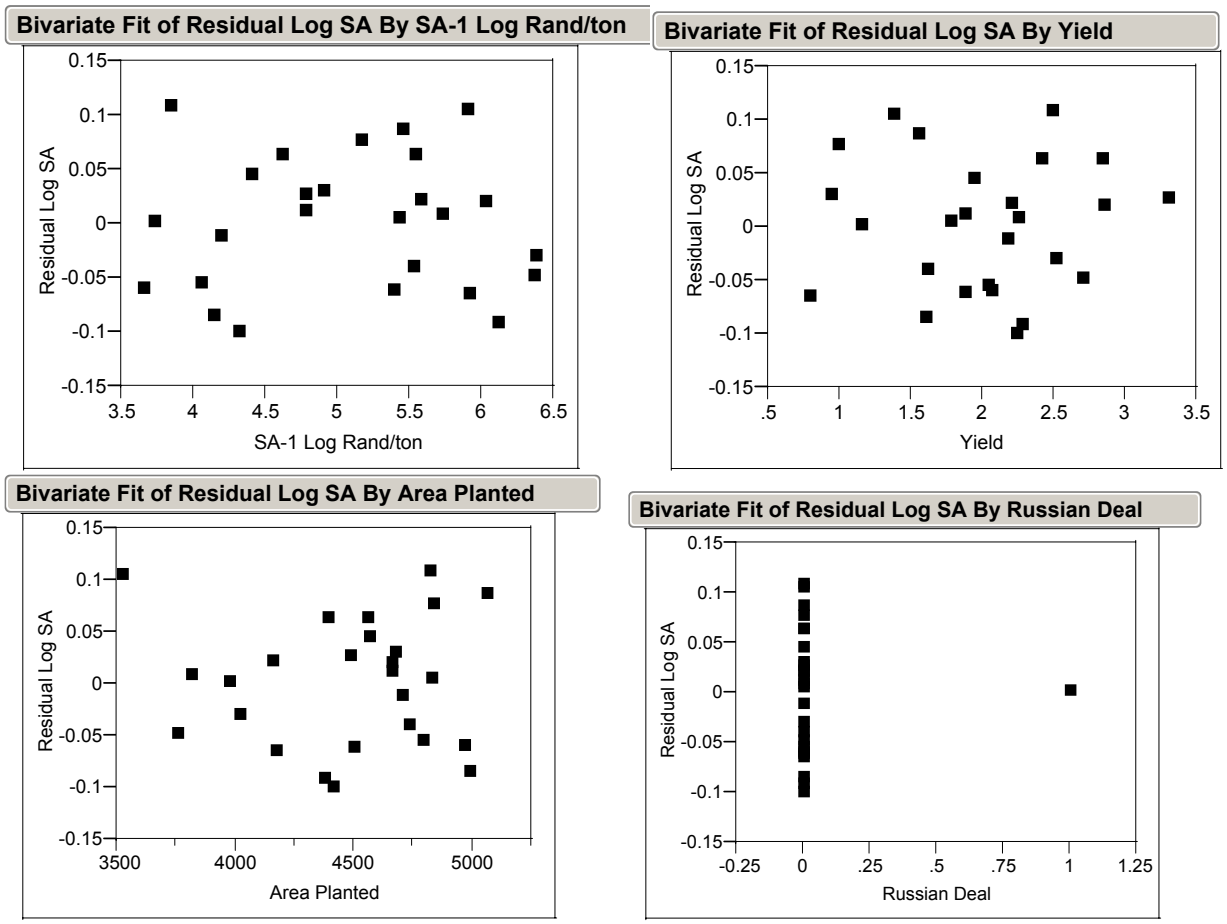
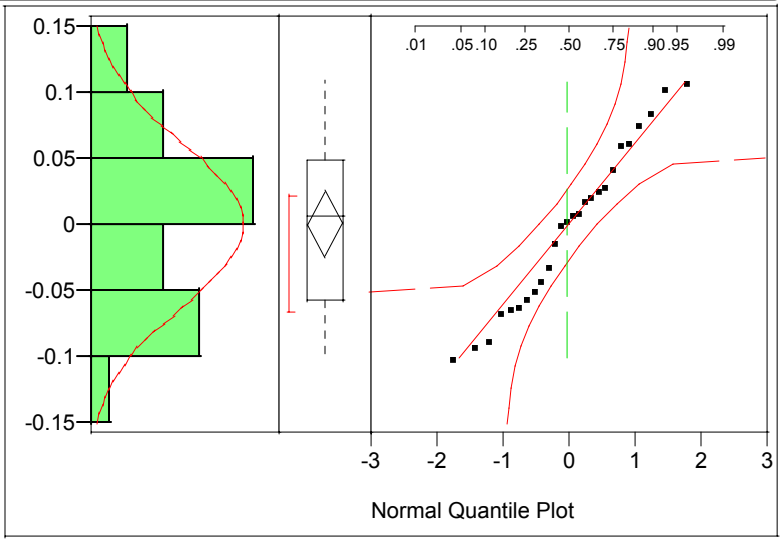


Figure 13. Residual vs. Predictor Plots

Distributions

Residual Log SA



Normal(8.5e-16,0.0614)

Quantiles

100.0%	maximum	0.1080
99.5%		0.1080
97.5%		0.1080
90.0%		0.0909
75.0%	quartile	0.0482
50.0%	median	0.0059
25.0%	quartile	-0.0572
10.0%		-0.0882
2.5%		-0.1006
0.5%		-0.1006
0.0%	minimum	-0.1006

Moments

Mean	8.54e-16
Std Dev	0.0613985
Std Err Mean	0.0120412
upper 95% Mean	0.0247994
lower 95% Mean	-0.024799
N	26

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	Mu	0.0000000	-0.024799	0.0247994
Dispersion	Sigma	0.0613985	0.048152	0.0847550

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.960609	0.4232

Figure 14. Normality of Residuals

Analysis of the studentized residuals show that only 1994 lies outside +/- 2. In the preliminary analysis, the studentized residual did not show any problems. In 1994, South Africa held its first democratic election, which in all likelihood is the cause for the borderline outlier. The outlier had minimal impact on the model, so the outlier is recognized but not corrected.

Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983
Studentized Resid	-1.00	0.00	1.77	-0.89	-1.39	-0.20	-1.61	0.68	0.97	0.44	0.16	0.49	1.26
Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Studentized Resid	-0.96	0.07	1.47	-0.65	0.99	0.33	0.13	-1.16	-1.47	0.33	2.02	-0.84	-0.51

Table 2. Studentized Residuals

Interpretation

At this point, the results validate the model represented. It is also important to interpret the results of the model. This model is complicated in that it contains both additive and multiplicative components. Equation 2 transforms Equation 1 to the South African producer price of maize using the exponential transformation.

$$p_{sa} = e^{(2.0608619+0.89375*\ln p_{sa-1}-0.132124*Yield-0.000254*Area-0.383306*Russian)} + e_t \qquad \textbf{Equation 2}$$

The coefficients of the model have the expected sign. The previous year's price is positive and the supply variables are negative. The supply variables should be negative, because as yield increases and the area planted increases, the production will increase and price will decrease. You will also notice that Russian Deal with the U.S. had a negative impact on South Africa.

A simple model would just use the lagged producer price as a predictor variable for the producer price. However, the significance level of each additional predictor variable, that is yield, area planted, and Russian Deal, is very significant and adds to the model.

It is also appropriate to include the confidence and prediction intervals in the analysis. Table 3 displays the predicted values of the South African producer price of maize as well as the confidence and prediction intervals. The values have been converted from natural logarithms using the exponential. The confidence intervals are on the mean and the prediction intervals are for a point estimate. Both intervals are narrow.

Year	Predicted Value	Lower 95% CI	Upper 95% CI	Lower 95% PI	Upper 95% PI
1971	44.310688	41.7798432	46.9948406	38.092026	51.5445692
1972	46.9900048	40.8791016	54.0144099	38.5868835	57.2230858
1973	51.861068	48.9774488	54.9144646	44.6102569	60.2904032
1974	66.6727087	63.6142187	69.878247	57.5571589	77.2319234
1975	72.5767392	69.0330695	76.3023159	62.5901509	84.1567404
1976	75.9143854	72.6454098	79.3304619	65.5951873	87.8569626
1977	90.2686075	85.7938551	94.9767496	77.8268446	104.699369
1978	97.446379	93.5227256	101.534645	84.2721035	112.680192
1979	111.492522	107.325082	115.821784	96.4984561	128.816387
1980	115.868278	108.41265	123.836635	99.2931428	135.210321
1981	134.117055	129.966558	138.400099	116.267597	154.706772
1982	169.985883	160.423831	180.11788	146.18139	197.666751
1983	203.934317	192.012253	216.596623	175.215457	237.360369
1984	242.227501	234.861375	249.824656	210.015155	279.38061
1985	232.65043	221.53061	244.328414	200.710394	269.673243
1986	231.137871	215.782638	247.585793	197.88015	269.985218
1987	266.402663	254.108354	279.291798	229.958541	308.622496
1988	250.726195	238.633267	263.431941	216.271638	290.669757
1989	300.103792	289.171052	311.449868	259.811037	346.645342
1990	368.836403	348.18135	390.716769	317.217635	428.854759
1991	485.824047	451.975536	522.207478	415.268274	568.367534
1992	453.65247	431.656141	476.769688	391.276752	525.971867
1993	359.883819	334.211793	387.527807	307.362762	421.379488
1994	523.359952	478.56951	572.342437	443.500698	617.599117
1995	624.394948	584.717213	666.765134	535.27027	728.359248
1996	610.745911	577.115444	646.336138	525.46939	709.861649

Table 3. Confidence and Prediction Intervals

Conclusion

Based on the results of this model, the hypothesis that internal factors are more influential in the price determination than external factors is proven. For the period of 1970 to 1996, South Africa used a simple method to determine price. The price of the previous year explains the vast majority of the variation in the current price. Addition variation in the current price could be explained by supply factors: yield and area planted. The ability of the Maize Board to regulate the price in South Africa over this period is quite interesting.

This analysis excluded the period of 1997 to 2001. The price determination for this period changed with the deregulation of the market. During the period with the Maize Board, the South African producer price of maize was more correlated to the previous year's price than the U.S. price. This changed in 1997. Figure 15 shows the relationship of the U.S. price and South African price. Additional research should evaluate the market from 1997 to the present. Using annual data, there are not enough observations, so it is suggested that monthly prices be analyzed.

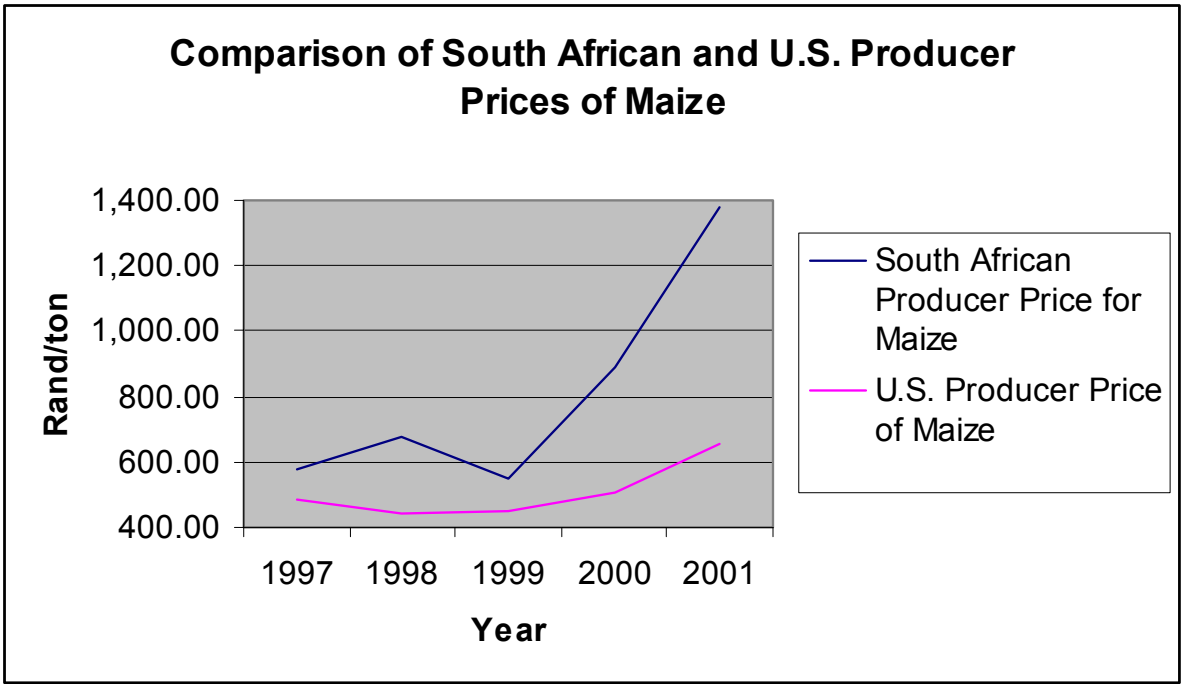


Figure 15. Relationship of U.S. price and South African price after Maize Board