

ASSESSING THE BENEFITS OF RESEARCH EXPENDITURES ON MAIZE PRODUCTION IN SOUTH AFRICA

R.F. Townsend¹, J. van Zyl² & C. Thirtle³

This paper focuses on assessing the benefits of research expenditures on maize production in South Africa. Both the production and supply function approaches are used to calculate elasticities of research expenditure on output and yield. Cointegration is used to establish long-run relationships between variables in these models. The lag structure of R&D expenditures on output is examined making use of the unrestricted, polynomial, beta and gamma distributions. The coefficients of these lag distributions were then used to calculate a rate of return to maize research expenditure, which was estimated as being between 28% and 39% per annum. These rates of return are high, mitigating in favour of more research expenditure rather than less.

RAMING VAN DIE VOORDELE VAN NAVORSINGSBESTEDINGS OP MIELIEPRODUKSIE IN SUID-AFRIKA

Hierdie referaat is gefokus op 'n raming van die voordele van navorsingsbestedings op mielieproduksie in Suid-Afrika. Beide die produksie- en aanbodfunksiebenaderings word gebruik om elastisiteite van navorsingsbesteding op produksie en opbrengs te bereken. Kointegrasie word gebruik om langtermyn verhoudings tussen veranderlikes in hierdie modelle te bepaal. Die sloerstruktuur van N & O bestedings op produksie word ondersoek deur die onbegrensde, polinomiale, beta- en gammaverdelings te gebruik. Die koëffisiënte van hierdie sloerverdelings is daarna gebruik om 'n opbrengskoers vir mieliënavorsingsbesteding, wat beraam is as tussen 28% en 39% per jaar, te bereken. Hierdie koerse is hoog en begunstig groter, eerder as kleiner navorsingbesteding.

1. INTRODUCTION

Agriculture's importance in the process of economic growth (van Zyl *et al*, 1988) highlights the role of sustained advances in farm production practices by improving the quality and quantity of farm products. Over the past several decades, maize has been the most important crop in South Africa, being both the major feed grain and the staple food for the majority of the South African

¹ Department of Agricultural Economics, Extension and Rural Development, University of Pretoria.

² Vice-Chancellor, and Principal, University of Pretoria.

³ Department of Agricultural Economics, University of Reading and Department of Agricultural Economics, Extension and Rural Development, University of Pretoria.

population. Maize accounts for about 40% of the value of crop production and about 15% of total agricultural production. In this context, investment in improved agricultural technology, particularly for maize, continues to be an important avenue for growth.

Over the past several decades, South Africa has been predominantly self-sufficient in food products and a major exporter of food grains. This was achieved under a regime of farmer support programmes which favoured large-scale commercial farmers. High production levels were encouraged by providing protection from foreign competition in the form of subsidies. Thus, the prices paid to farmers were frequently higher than world prices (Kirsten & Van Zyl, 1996). These policies created distortions, including the planting of maize in areas of marginal land, which led to environmental degradation and encouraged unsustainable farming practices Brand *et al.*, 1992).

Since the mid-1980s, there has been widespread deregulation of marketing and liberalisation of price controls in the agricultural sector. As a result of this movement to a more market based environment, the real price of maize has declined by 50% since 1983, which has caused the area planted to maize to decline, particularly since 1986/87. Within the changing environment of falling real producer prices and rising population, the successful introduction of new technology is a major factor in maintaining profits and for maintaining food self-sufficiency. Maize research and development is largely funded by the public sector and increased pressure on the Government budget has emphasised the need for prior assessment of the potential benefits of research in order to secure government funding.

This paper will focus on assessing the benefits of research expenditures on maize production in South Africa. The next section will discuss the alternative methodologies used to assess research benefits. Tests for long-run relationships between maize output and research expenditure will be conducted in section three. Having determined these long run relationships, section four will examine the lag structure between research expenditure and output. These lag structures will then be used to calculate the rate of return to maize research in the final section, before the conclusion.

2. METHODS OF MEASURING RETURNS TO RESEARCH

The relationship between R&D expenditures and technology-based output or productivity growth in agriculture has attracted the attention of both historians and economists at least since Griliches (1958). The progress to date is well documented in Alston *et al.*, (1995) and the literature on the returns to R&D is

surveyed by Echeverria (1990). The majority of empirical rate of return (ROR) models reported are based on either *economic surplus* calculations, following Griliches (1958), or on *econometric estimation* of the output elasticity of R&D, derived from the production relationship (also pioneered by Griliches, in the early 1960s). This paper will focus on the econometric approach using both the supply and the production function:

The production function approach

The production function for maize in this study was specified as

$$YIELD = f(EXT, RD, RAIN) \quad (1)$$

where YIELD is tons per hectare, RD is the research expenditure, EXT is a two-year moving average of expenditures and RAIN is regional rainfall. Conventional inputs such as fertiliser may also be expected to affect yields, but crop-specific fertiliser data was not available. Although ROR calculations are made on the basis of estimating equations that do not include conventional inputs (see Akgungor *et al.*, 1996 for a recent example), Alston *et al.* (1995:107) correctly argue that conventional inputs should be included. This is a potentially serious problem since fertiliser, R&D, extension expenditures and yields are all positively correlated. Thus, omitting fertiliser could bias the R&D elasticity upwards and inflate the ROR. In this case, extension and fertiliser use are highly correlated in South Africa so including both variables in the same equation may cause collinearity problems. However, extension may be used as an instrumental variable for fertilizer.

The supply response estimation

For the general case, the supply function for maize output, $Output_m$, can be expressed as

$$Output_m = f(P_m, P_s, P_{sun}, P_{nuts}, P_{fert}, P_{lab}, RISK, RD, RAIN) \quad (2)$$

where P_m is the price of maize. The competing crops in the production of maize are sorghum, P_s , sunflowers, P_{sun} , and groundnuts, P_{nuts} , with fertiliser, P_{fert} , and labour, P_{lab} , the key production costs. These factors affect any movement along the supply curve. The weather, RAIN, and R&D expenditures are included to capture shifts in the supply curve. The R&D expenditures included in this study are those of the Grain Crops Institute which covers fields such as plant breeding, cultivar evaluation, grain quality, plant physiology, tillage, plant nutrition, weed control, plant pathology, entomology, nematology and yield performance (Van Zyl & Sartorius von Bach, 1994).

In order to capture the farmers' response to the price variations, a price risk variable (RISK) is included. The hypothesis is that competitive, risk-averse farmers produce less when there is price uncertainty. Only price risk was included in the analysis as some yield risk can be borne by insurance companies, reducing the impact on output. There has been considerable debate on the use of risk as a factor in econometric studies of the supply of agricultural commodities (see Traill, 1978; Gallagher, 1978, 1974; Just, 1974; Chembezi, 1991). The methods used vary from simple measures of instability to complex variables requiring complex estimation procedures. The method used by Ryan (1974) and adapted by Chembezi (1991) is used in this analysis.

Under certain assumptions about the decision-maker, Ryan (1974) expressed output as a function of the interaction terms between expected prices and their variances:

$$x = f \left[\frac{\sigma_{ij}^2}{P_i^*}, \frac{\sigma_{ij}^*}{P_i^* \sigma_j^2}, \frac{\sigma_i^2 \sigma_j^2}{P_j^* \sigma_{ij}}, \frac{\sigma_{ij}}{P_{ij}} \right] \tag{3}$$

where x_i is the output, P_i^* and P_j^* are the expected producer prices, σ_i^2 and σ_j^2 are the variances of the i th and j th products and σ_{ij}^* is the covariance of products i and j . The first term shows how the effect of crop price variance on output is modified by the level of the expected crop price. The second term shows that the effect of the covariance between crop prices and the acreage planted is modified by the level of crop price and the variance of the price of the competing crop price. The other two terms are hard to interpret while risk and the expected price are assumed to be additively separable. To generate a price risk variable Gallagher's (1978) modification of Ryan's (1974) formulation, is a proxy for the first term in equation (2) and ignores the other terms.

$$RISK_t = \frac{(P_{t-1} - \bar{P})^2}{\bar{P}} \tag{4}$$

\bar{P} is expressed using Fisher's distributed lag weights instead of equal weights as suggested by Gallagher (1978) where $\bar{P} = 0.50P_{t-2} + 0.33P_{t-3} + 0.17P_{t-4}$.

3. TESTING FOR LONG-RUN RELATIONSHIPS BETWEEN THE VARIABLES

The data used in the analysis is mainly from the *Abstract of Agricultural Statistics* (Republic of South Africa, 1972, 1996), unpublished information from the Department of Agriculture and the Weather Bureau. Long-run relationships between the variables in equation (1) and (2) can be determined using cointegration tests. A pre-condition for cointegration is that the variables have the same statistical properties which can be tested using the Dickey-Fuller test, (1981). All the variables appear to have the same properties, being integrated of order one, $I(1)$. The Dickey-Fuller ϕ_2 tests show that, in the levels, all of the variables have a unit root, with no drift or trend. Having established the order of integration, the next stage tests for cointegration between the variables to establish the long-run relationships. Many of the variables described in equation 2 were found to have no explanatory power in these regressions and were eliminated. Three tests are used to test for cointegration, namely the Dickey-Fuller test, the CRDW and the Johansen maximum likelihood approach. Table 1 first reports the results and critical values (in brackets).

Table 1: Cointegration tests

Equation	Dickey-Fuller Test	CRDW (0.68)	Johansen Tests	
			Eigen value	Trace
YIELD = EXT+RD+MRRAIN	-5.04 (-4.36)	1.48	41.29 (41.29) 15.67 (8.34)	52.83 (34.91) 19.96 (11.53)
Output _m = P _{mf} +RISK+P _{sf} +RD+MRRAIN	-6.04 (-4.77)	1.89	44.92 (31.32) 26.97 (27.14)	98.48 (70.59) 53.55 (48.28)

Values in brackets are the 5% significance levels.

In Table 1, YIELD is tons/ha of maize, EXT is the extension expenditure, P_{mf} and P_{sf} are ratios of the producer price of yellow maize to fertiliser and sorghum to fertiliser price respectively; Output_m is maize output (1000 tons); RD is the research expenditure for the summer grain complex; RISK is the risk variable; MRRAIN is average rainfall for the maize producing areas from September to February. All the tests show the variables to be cointegrated and hence there exists a long run relationship between the variables.

4. LAGGED EFFECTS OF R&D EXPENDITURES ON YIELD AND OUTPUT SUPPLY

In order to capture lagged effects of R&D on output and avoid the collinearity problem of the unrestricted lag model, a common approach is to use an Almon polynomial lag Evenson, 1967; Lu *et al.*, 1978, 1979, Knutson & Tweeton, 1979;

Doyle & Ridout, 1985; and Thirtle & Bottomely, 1988, 1989). The polynomial form is popular due to its empirical simplicity, providing a smooth and feasible form. However, the specification may require restrictions and the validity of these has been questioned, particularly end point restrictions (Hallam, 1990), and there are suggestions that these models may lead to biased estimates of the output effects of research spending. To avoid these biased results, less restrictive forms such as the beta and gamma distributions (derived from the Pearson representation) as well as the unrestricted model can be used. The Akaike Information Criteria (AIC) and the Schwarz Criteria (SC) are used to determine not only the lag length but the degree of polynomial in the model.

Production function approach

First, the lag lengths are determined using the unrestricted model for maize according to the Akaike Information Criteria (AIC) and the Schwarz Criteria (SC). The lag was found to be 6 years and the sum of the lag coefficients was 0.97. The Hatanaka & Wallace procedure (1980), employed by Silver and Wallace (1980), Hallam, (1990) and Khatri (1994), using lower order moments, where the zero order of the moment is equivalent to the sum of the lag coefficients, gave the same coefficients with t-statistics of 4.46. The coefficients of the unrestricted lag terms change sign frequently (due to collinearity) as shown in the first columns of Table 2. This can not be justified theoretically and the sign changes make internal rate of return calculations impossible. Therefore, a structure must be imposed.

The polynomial lag (PDL) model was estimated with no restrictions, as well as near end, far end and both end points restricted to equal zero. These restrictions were applied to second, third and fourth order polynomials for a range of lag lengths, using the AIC and SC as the model selection criteria. The model chosen based on these tests was a second degree polynomial with a lag of 12 years. This lag length is much greater than those suggested by the unrestricted model. Indeed, the unrestricted lag length appears to correspond with the peak effect of R&D on yields, rather than the total lag length. Thus, at least for this data, determination of the lag length using the unrestricted model as suggested by Hallam (1990) gives results that are at odds with the outcomes of the Andrews & Fair (1992) PDL procedure.

The gamma and beta distributions were also fitted to see if the distributions were skewed, rather than symmetric. These results are shown in Table 2 and Figure 1. The peak effect of R&D on maize yields occurs after six years, for all

Table 2: Lag structure for R&D in the yield and supply equations

Variable	Maize Yield Equation				Maize Supply Equation	
	Unrestricted	PDL	GAMMA	BETA	Unrestricted	PDL
Constant	-2.641 (-1.0)	-3.193 (-1.3)	-4.562 (-1.6)	-4.842 (-1.7)	3.747 (1.6)	3.044 (1.8)
P _{mf}	-	-	-	-	0.653 (1.2)	0.842 (2.1)
RISK	-	-	-	-	-0.039 (-1.6)	-0.035 (-1.7)
P _{sf}	-	-	-	-	-0.813 (-2.3)	-0.785 (-2.6)
Rainfall	0.615 (2.5)	0.770 (3.84)	0.790 (3.7)	0.689 (3.0)	0.669 (2.4)	0.786 (3.7)
Extension	0.610 (1.2)	0.514 (1.39)	0.746 (1.7)	0.879 (1.9)	-	-
R&D	-0.546 (-0.9)	0.029 (4.7)	0.000 (0.0)	0.000 (0.0)	0.276 (0.4)	0.026 (1.8)
R&D _{t-1}	0.128 (0.2)	0.055 (4.7)	0.001 (0.2)	0.001 (0.3)	-0.236 (-0.3)	0.047 (1.8)
R&D _{t-2}	1.442 (1.7)	0.076 (4.7)	0.022 (0.3)	0.034 (0.5)	0.726 (0.8)	0.061 (1.8)
R&D _{t-3}	-0.221 (-0.3)	0.092 (4.7)	0.071 (0.8)	0.082 (1.1)	-0.169 (-0.2)	0.070 (1.8)
R&D _{t-4}	-0.955 (-1.3)	0.103 (4.7)	0.125 (2.0)	0.126 (2.3)	-0.374 (-0.4)	0.073 (1.8)
R&D _{t-5}	-0.113 (-0.2)	0.110 (4.7)	0.162 (2.3)	0.160 (2.4)	-0.258 (-0.3)	0.070 (1.8)
R&D _{t-6}	1.250 (2.4)	0.113 (4.7)	0.171 (1.9)	0.176 (2.0)	0.746 (0.8)	0.061 (1.8)
R&D _{t-7}		0.110 (4.7)	0.157 (1.9)	0.170 (1.9)	0.621 (0.7)	0.047 (1.8)
R&D _{t-8}		0.103 (4.7)	0.130 (2.2)	0.150 (2.2)	-0.838 (-1.2)	0.026 (1.8)
R&D _{t-9}		0.092 (4.7)	0.100 (2.2)	0.106 (1.9)		
R&D _{t-10}		0.076 (4.7)	0.072 (1.6)	0.064 (1.0)		
R&D _{t-11}		0.055 (4.7)	0.050 (1.0)	0.027 (0.6)		
R&D _{t-12}		0.030 (4.7)	0.033 (0.7)	0.001 (0.3)		
SUM (R&D)	0.969	1.045	1.093	1.107	0.494	0.479

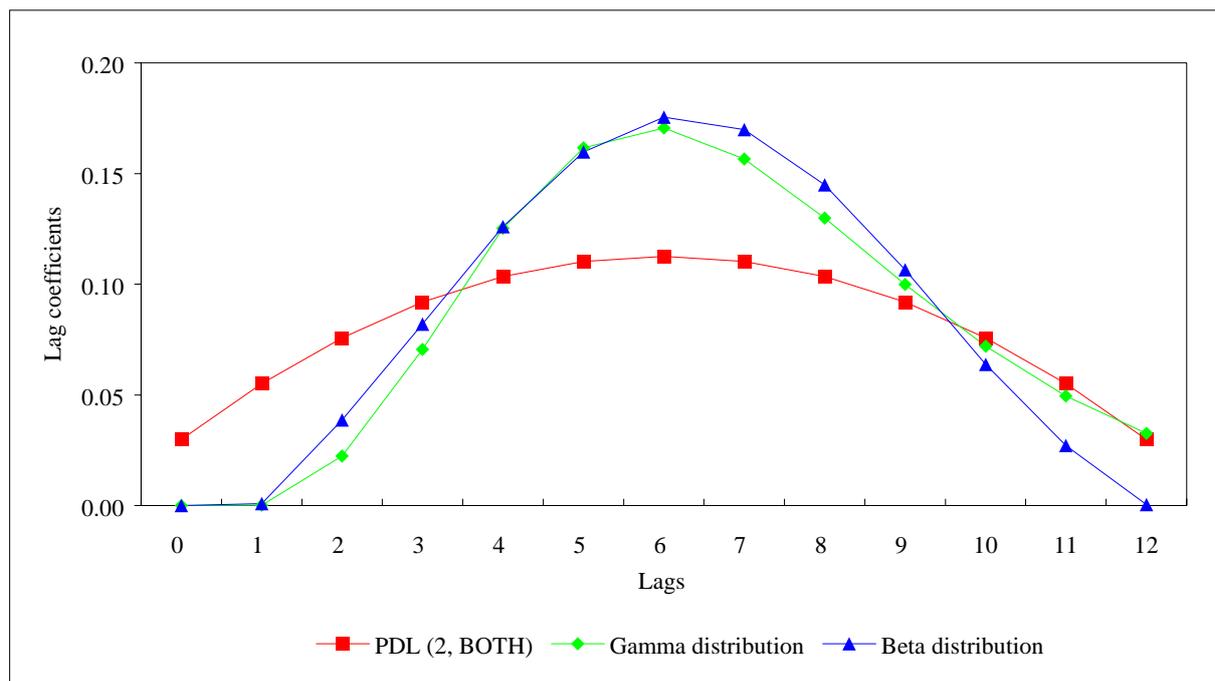


Figure 1: Second degree polynomial, gamma and beta distribution of R&D lags on maize yield.

three models. The beta and gamma distributions show very similar lag structures, with a possible lead of one year and both are very nearly symmetric. For both distributions, the most significant lags are from four to nine years, as the t-statistics show. The sums of the lag coefficients are 1.045, 1.093 and 1.107 for the polynomial, gamma and beta lags respectively. All are slightly higher than the 0.97 derived from the unrestricted lag, but the range is not great enough to substantially affect the ROR results. This is just as well, since in this case, where the lag structures appear to be symmetric, there is little to choose between the models and the two criteria often fail to agree as to the preferred model. Thus, the much-maligned second degree PDL, with end point restrictions, is in this case as good as any more complex model.

Output supply approach

Using the same procedure, the lag length of R&D expenditure in the supply model was found to be 8 years. The lag shape results for the output supply functions were less interesting because the lag terms for the gamma and exponential distributions were all insignificant. Thus, only the results for the unrestricted model and the PDL with end point restrictions results are reported in Table 2.

The lag structure identifies the effects of changes in R&D expenditures on the output and yield of maize. To calculate a rate of return (ROR) value, the elasticities have to be converted to value marginal products. Using the procedure in Thirtle & Bottomley (1988), the value marginal product of R&D in period $t-i$ can then be written as

$$VMP_{t-i} = \frac{\Delta VALUE_t}{\Delta RD_{t-i}} = \beta_i \frac{\overline{OUTPUT}}{\overline{RD_{t-i}}} \cdot \frac{\Delta VALUE_t}{\Delta OUTPUT_t} \quad (5)$$

where $OUTPUT/RD_{t-i}$ is an average and $\Delta VALUE_t/\Delta OUTPUT_t$ is calculated as the average of the last five years minus the average for the first five years⁴, for both variables. Thus, these are constants, but β_i varies over the lag period, giving a series of marginal returns resulting from a unit change in R&D expenditure. The value of output, $\Delta VALUE_t/\Delta OUTPUT_t$ is the geometric mean calculated using the value of output at constant 1950 prices. Similarly, $OUTPUT/RD_{t-i}$ is a constant-price geometric average. The marginal internal rate of return (MIRR) is calculated from equation (6),

$$\sum_{i=1}^n \frac{VMP_{t-i}}{(1+r)^i} - 1 = 0 \quad (6)$$

where n is the lag length, by solving for r . To estimate a ROR from the yield equation output can be substituted for yield in equation (5).

The MIRRs calculated for maize, using different lag structures, are reported in Table 3. Note that at the crop level there are no net output measures, such as net farm income. The value used is the gross value of output, without the value of inputs being netted out. This can be done if there is gross and net margin information that is representative of the whole country and consistent over the period. Net returns, calculated in this way would be lower, but still substantial. Table 3 shows the MIRR results. The ROR for maize research range from 28% to 39%.

The effects of the alternative lag structures on the MIRR for maize is not very great. For maize, the PDL gives a lower estimate than the gamma and beta models because of the higher elasticities in the middle period.

⁴ The averaging is necessary because of the fluctuations.

Table 3: Marginal internal rates of return to R&D expenditures on maize

Lag Structure	Marginal Internal Rate of Return (MIRR)
Second-order PDL, end point restrictions (Output supply equation)	30.32
Second-order PDL, end point restrictions (Yield equation)	28.84
Gamma distribution (Yield equation)	36.87
Beta distribution (Yield equation)	39.84

5. CONCLUSION

The rate of return on maize research and development was calculated as being 28% and 39%. This result confirms the importance of allocating resources to investment in research and development to increase yields and output. In an environment of falling real producer prices and rising population the successful introduction of new technology is a major factor in maintaining profits and for maintaining food self-sufficiency.

Increasingly, there are pressures for productivity-enhancing research to be funded by levies on producers. In South Africa, this may be possible for the large-scale commercial farmers, who have been the chief beneficiaries of public R&D in the past. This bias towards commercial production and lack of emphasis on technology development for the communal areas needs to change. The farmers in resource poor areas, with lower soil quality and rainfall, as well as low input use, need different technologies to suit their own needs. Thus, the focus of technology development needs to be changed so that the R&D investment for the communal areas reflects both the numbers involved and the abject poverty of a considerable proportion of the population.

At the same time, research policy must maintain a balance between the need for assisting the communal areas and maintaining the superior quality of output and higher productivity of the growth areas of the commercial sector. There are also significant positive spillovers from technical change in these highly productive regions which can benefit the poor people in marginal areas. The poor also benefit through lower food prices, increased employment and higher wages. These effects need to be balanced against the benefits to be gained by re-targeting research expenditures towards resource-poor farmers. The problems of transforming traditional agriculture are sufficiently great that the returns

may be low for some time. However, research directed towards those technologies that enhance moisture conservation and the efficient use of rain, such as conservation tillage, crop rotation, clean fallow, collection of water run-off and improved weed control, are potentially highly effective in marginal environments.

REFERENCES

AKGUNGOR, S., MAKANDA, D., OEHMKE, J., MYERS, R. & CHOE, Y. (1996). A dynamic analysis of Kenyan wheat research and rate of return, pp. 1-31. In *Conference Proceedings, Contributed Papers, published by the Conference Secretariat of the Global Agricultural Science Policy for the Twenty-first Century*, Melbourne, Australia, August.

ALSTON, J., CRAIG, B. & PARDEY, P. (1996) Research lags and research returns, pp.333-366, in *Conference Proceedings, Invited Papers, published by the Conference Secretariat of the Global Agricultural Science Policy for the Twenty-first Century*, Melbourne, Australia, August 1996.

ANDREWS, D. W. & FAIR, R.C. (1992). Estimation of polynomial distributed lags and leads with end point constraints. *Journal of Econometrics*, 53:123-139.

BRAND, S., CHRISTODOULOU, N., VAN ROOYEN, C.J. & VINK, N. (1992). Agriculture and redistribution. In R. Schrire (eds), *Wealth or poverty? Critical choices for South Africa*. Cape Town : Oxford University Press.

CHEMBEZI, D. (1991). Acreage response to price and risk : The case of tobacco in Malawi. *European Review of Agricultural Economics*, 18:85-102.

DOYLE, C.J. & RIDOUT, M.S. (1985). The impact of scientific research on UK agricultural productivity. *Research Policy*, 14:251-276.

DICKEY, D.A. & FULLER, W.A. (1981). Likelihood ratio statistics for autoregressive time series with unit roots. *Econometrica*, 49:1057-72.

ECHEVERRIA, R.G. (1990) (ed) *Methods for diagnosing research system constraints and assessing the impact of agricultural research*, Vol. 2. The Hague : ISNAR.

EVENSON, R.E. (1967). The contribution of agricultural research to production. *Journal of Farm Economics*, 49(5):1415-25.

GALLAGHER, P. (1978). The effectiveness of price support policies - some

evidence for US corn acreage response. *Agricultural Economics Research*, 30(4):8-12.

GRILICHES, Z. (1958). Research costs and social returns : Hybrid corn and related innovations. *Journal of Political Economy*, 66:419-31

HALLAM, D. (1990). Agricultural research expenditures and agricultural productivity change. *Journal of Agricultural Economics*, 41:434-9.

HATANAKA, M. & WALLACE, T.D. (1980). Multicollinearity and the estimation of low-order moments in stable lag distributions. In J. Kmenta and J.B. Ramsey. (Eds). *Evaluation of econometric models*. New York : Academic Press.

JUST, R.E. (1974). An investigation of risk in farmers decisions. *American Journal of Agricultural Economics*, 56:14-25.

KHATRI, Y.J. (1994). *Technical change and the returns to research in UK agriculture 1953-1990*. Unpublished PhD thesis, University of Reading.

KIRSTEN, J. & VAN ZYL, J. (1996). The contemporary agricultural policy environment: undoing the legacy of the past. In J. van Zyl, J. Kirsten and H.P. Binswanger. (Eds), *Agricultural land reform in South Africa*. Cape Town : Oxford University Press.

KNUTSON, M. & TWEETEN, L.G. (1979). Towards an optimal rate of growth in agricultural production research and extension. *American Journal of Agricultural Economics*, 61:70-6.

REPUBLIC OF SOUTH AFRICA. (1972, 1996). *Abstract of agricultural statistics*. Directorate of Agricultural Economic Trends, Department of Agriculture and Water Supply, Pretoria : Government Printer.

RYAN, J.J. (1974). Supply response to risk: The case of US pinto beans. *Journal of Agricultural Economics*, 2:35-43.

THIRTLE, C. & BOTTOMLEY, P. (1988). Is publicly funded agricultural research excessive? *Journal of Agricultural Economics*, 39:99-111.

THIRTLE, C. & BOTTOMLEY, P. (1989). The rate of return to public sector R&D in the United Kingdom, 1965-1980. *Applied Economics*, 21:1063-1086.

TRAILL, B. (1978). Risk variables in econometric supply models. *Journal of Agricultural Economics*, 29:53-61.

VAN ZYL, J. & SARTORIUS VON BACH, H.J. (1994). *Statistical brief on the National Agricultural Research System of South Africa*. Statistical Brief, ISNAR Indicator Series Project: Phase II.

VAN ZYL, J. & NEL, H.J.G. (1988). The role of the maize industry in the South African Economy. *Agrekon*, 27(3):10-16.