Examination of Alternative Heteroscedastic Error Structures Using Experimental Data

James W. Mjelde, Oral Capps, Jr. and Ronald C. Griffin*

Abstract

Impacts of alternative specifications for heteroscedastic error structures are examined by estimating various production functions for corn in Central Texas. Production- and profit-maximizing levels of inputs and the shape of the profit equation obtained from models not corrected for heteroscedasticity differed from those obtained from models corrected for heteroscedasticity. Using the profit-maximizing input levels for each production function gave essentially the same estimated yield and profit, regardless of the specification for heteroscedasticity employed. Differences of up to one-quarter to one-third are noted, however, in the amount of profit-maximizing levels of inputs used, depending on the heteroscedasticity correction.

Key words: corn, heteroscedasticity, production functions

Evaluating crop yield response to inputs is fundamental to studies modeling producer reactions to changing environmental and economic conditions. Estimation of the dependence of yield on inputs usually involves employing ordinary least squares (OLS) regression. One potential problem in using OLS is the presence of heteroscedastic error terms. Heteroscedasticity exists when the variance of the error terms is not constant among observations. Statistical consequences of heteroscedasticity in OLS estimations are: 1) the estimates of the regression model are unbiased, but asymptotically inefficient, and (2) the estimates of the variances of the regression coefficients are biased (Maddala). Yang, Koo, and Wilson contend that ". . . heteroscedasticity has received less attention and frequently has been handled inadequately in empirical analysis" (p. 103). The present study addresses this issue by examining alternative techniques to correct for heteroscedasticity. We also consider the influence these techniques have on production- and profit-maximizing levels of inputs.

Generalized least squares (GLS) is normally used to improve asymptotic efficiency of the parameter estimates when heteroscedasticity is present. In practice, a major difficulty lies in knowledge of the particular form to employ to model heteroscedastic error terms. Many different functional forms can be used to form the GLS estimator. Judge *et al.* (1985) contend "there is no well-established 'best way' to model heteroscedasticity" (p. 454). Further, Judge *et al.* (1985) hypothesize that the choice among heteroscedastic structures is not likely to be important. The choice of which structure to use may be based on estimation convenience, because *a priori* no "best way" exists to model heteroscedasticity. In this light, the objective of this study is to empirically consider the impacts of

*The authors are associate professor and professors in the Department of Agricultural Economics, Texas A&M University. This work was funded by Project Numbers 6507 and 7161 of the Texas Agricultural Experiment Station. This manuscript is Technical Article No. 31873 of the Texas Agricultural Experiment Station.

J. Agr. and Applied Econ. 27 (1), July, 1995: 197-211

alternative specifications for heteroscedastic error structures in estimating corn production functions for Central Texas. Emphasis is placed upon whether the choice of specification influences optimal input recommendations made to corn producers.

Heteroscedastic Models

To examine heteroscedasticity, consider the following general linear model

$$y_t = f(x_t, \beta) + v_t \tag{1}$$

where y_t corresponds to output, x_t refers to a vector of inputs, β is a vector of unknown parameters to be estimated, and v_t is a stochastic disturbance term. Heteroscedasticity exists when

$$E(v_t) = 0, \textit{ and}$$
$$E(v_t, v_s) = \begin{cases} \sigma_t^2 \textit{ for } t = s, \\ 0 \textit{ for } t \neq s. \end{cases} \tag{2}$$

Equation (2) implies that the variance of the disturbance term may vary from observation to observation Judge *et al.* (1982). If σ_t^2 is a constant (that is $\sigma_1^2 = \sigma_2^2 = \ldots \sigma_T^2$) then the disturbance is homoscedastic and this assumption of ordinary least square (OLS) is not violated. Usually, the σ_t^2 are not known; therefore some method must be used to obtain estimated variances. These estimated variances are then used in a generalized least squares estimator. Judge *et al.* (1980) notes that "because there are T unknown variances and T observations, it is unlikely that we could obtain a reasonable variance estimator without some further assumptions that reduces the number of unknown parameters" (p. 416). Usually, it is assumed that each σ_t^2 is a function of S explanatory variables. These variables could be anything, but in practice the variables are selected to be a subset of the exogenous variables. With this assumption, the equation to estimate variances is:

$$\sigma_t^2 = h(x_t, \gamma) + e_t \tag{3}$$

where γ is a vector of parameters to be estimated.

Mechanically, the estimation procedure works as follows. First, the residuals (denoted by v) from the OLS regression of y_t on $f(x_t, \beta)$ are obtained. Second, OLS is used in the regression of some form of v_t on $h(x_t, \gamma)$ to obtain consistent estimates of γ. The form of v_t used depends on the heteroscedasticity correction being employed. In all cases, this form is an estimator for the variance or standard deviation of output. Third, GLS, or perhaps more properly weighted least squares, is used to obtain consistent and asymptotically efficient estimates of β. In this stage of the estimation procedure, both y_t and $f(x_t, \beta)$ are weighted by the inverse of the standard deviation of $h(x_t, \hat{\gamma})$ where $\hat{\gamma}$ are estimated parameters.

Given (1), the problem becomes not only choosing an appropriate functional form for $f(x_t, \beta)$ but also an appropriate specification for $h(x_t, \gamma)$. Our primary objective is to investigate whether the choice of form for $h(x_t,\gamma)$ matters for input recommendations. Alternative functional forms for $f(x_t, \beta)$ used are the quadratic, translog, and square root production functions. Six different specifications for estimation of $h(x_t, \gamma)$ are examined for each of the three production function specifications. Specification of $h(x_t,\gamma)$ requires both a functional form for $h(x_t, \gamma)$ and an estimator for the variance of output. A form of the estimator for the variance of output becomes the dependent variable in the OLS regression to obtain $\hat{\gamma}$.

Previous Specifications for the Variance of Output

Different tests for heteroscedasticity have been employed previously. These tests generally differ in two ways: (1) the functional form of $h(x_t, \gamma)$ and (2) the estimates of the variance of output. Several tests and previous studies are discussed here. For a more thorough discussion of heteroscedasticity see Carroll and Ruppert.

Hildreth and Houck, as well as Amemiya (1977), specified $h(x_t, \gamma)$ to be a linear function of the exogenous variables x_t. Mathematically, this specification is:

$$h(x_t, \gamma) = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \ldots + \gamma_k x_{kt}. \tag{4}$$

The auxiliary regression used to determine estimates of the γ_j 's is $v_t^2 = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \ldots + \gamma_k x_{kt}$, where v_t corresponds to the residuals for the regression of y_t on $f(x_t, \beta)$. This regression corresponds to the Breusch-Pagan-Godfrey test of heteroscedasticity (Breusch and Pagan). With this specification, the weights in the GLS procedure are

$$\left[\hat{\gamma}_0 + \sum_{j=1}^{k} \hat{\gamma}_j x_{jt} \right]^{-1/2}. \tag{5}$$

Just and Pope (1978, 1979) and Buccola and McCarl employed a form for $h(x_t, \gamma)$ resembling a Cobb-Douglas relationship, that is,

$$h(x_t,\gamma) = \gamma_0 x_{1t}^{\gamma_1} x_{2t}^{\gamma_2} \ldots x_{kt}^{\gamma_k}. \tag{6}$$

The auxiliary regression used to obtain estimates of the γ_j's is $\ln v_t^2 = \ln \gamma_0 + \gamma_1 \ln x_{1t} + \gamma_2 \ln x_{2t} + \ldots + \gamma_k \ln x_{kt}$. Just and Pope (1979) use, however, $\ln |v_t|$ in lieu of $\ln v_t^2$ as an estimate for the standard deviation of output. Given that $\ln v_t^2 = 2 \ln|v_t|$, the use of either $\ln v_t^2$ or $\ln|v_t|$ as the dependent variable in this auxiliary relationship to obtain estimates of $\hat{\gamma}$ is equivalent (Mjelde, Griffin, and Capps). This regression corresponds to the Park-Glejser test of heteroscedasticity (Park; Glejser). With this specification, the weights used in the GLS estimation are

$$\hat{\gamma}_0^{-1/2} \left[\prod_{j=1}^{k} x_{jt}^{\hat{\gamma}_j/2} \right]^{-1}. \tag{7}$$

Additionally, the logarithm of $h(x_t, \gamma)$ may be a linear function of the exogenous variables. Judge *et al.* (1985) term this form multiplicative heteroscedasticity, given by

$$h(x_t,\gamma) = \exp\left[\gamma_0 + \sum_{j=1}^{k} \gamma_j x_{jt} \right] \tag{8}$$

This specification was employed by Harvey and is popularly known as the Harvey test. The auxiliary regression used to derive estimates of the γ_j's is $\ln v_t^2 = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \ldots + \gamma_k x_{kt}$. With this specification, the weights used in the GLS procedure correspond to

$$\left\{ \exp\left[\hat{\gamma}_0 + \sum_{j=1}^{k} \hat{\gamma}_j x_{jt} \right] \right\}^{-1/2}. \tag{9}$$

Testing for heteroscedasticity in the various formulations is similar. The null hypothesis is that all coefficients, except the intercept, associated with $h(x_t, \hat{\gamma})$ are jointly equal to zero. A standard F-test is used to test for the existence of heteroscedasticity. Rejecting the null hypothesis implies the presence of heteroscedasticity.

Caution should be used when interpreting R^2 obtained from GLS estimation. These R^2 measures are based on transformed data; that is the data have been used three times: once in obtaining the estimates for the variance of output, again in obtaining estimates for γ, and finally in obtaining estimates for β. To circumvent this problem, the following procedure is employed to obtain goodness-of-fit measures. For each observation, estimated yields are obtained using the GLS estimates for the production function. Simple correlation coefficients between the actual and predicted yields are then calculated. The square of these correlation coefficients are then adjusted for degrees of freedom.

Heteroscedastic Specifications Examined

Based on the previous studies, different heteroscedastic error structures are examined for each of the three production functional forms. As noted earlier, these structures differ in assumptions on the form of $h(x_t, \gamma)$ and the estimator for the variance of output. Alternative specifications for $h(x_t, \gamma)$ and the variance of output are given in table 1. The exogenous variables in $h(x_t, \gamma)$ are precisely the same as those in $f(x_t, \beta)$.

Table 1. Alternative Estimated Heteroscedastic Error Models

Model Number	Estimator of the Variance of Output[a]	Specification of $h(x_t, \gamma)$	Popular Name
One	$\ln(v_t^2)$	$\gamma_0 \; x_{1t}^{\gamma_1} \; x_{2t}^{\gamma_2} \; . . . \; x_{kt}^{\gamma_k}$	(Park - Glejser)
Two	v_t^2	$\gamma_0 \; x_{1t}^{\gamma_1} \; x_{2t}^{\gamma_2} \; . . . \; x_{kt}^{\gamma_k}$	
Three	$\lvert v_t \rvert$	$\gamma_0 \; x_{1t}^{\gamma_1} \; x_{2t}^{\gamma_2} \; . . . \; x_{kt}^{\gamma_k}$	
Four	$\ln(v_t^2)$	$\gamma_0 + \sum_{j=1}^{k} \gamma_j x_{jt}$	(Harvey)
Five	v_t^2	$\gamma_0 + \sum_{j=1}^{k} \gamma_j x_{jt}$	(Breusch-Pagan-Godfrey)
Six	$\lvert v_t \rvert$	$\gamma_0 + \sum_{j=1}^{k} \gamma_j x_{jt}$	

a Dependent variable used in the auxiliary regression equation to estimate the variance of output.

Data

Data covering a five-year period, 1984-88, from seven field experimental plot studies for corn are used to investigate the effect of various heteroscedastic specifications. The experiments were conducted at two locations in Texas, Brazos River Bottom Research Farm and Stiles Research Farm, which are located approximately 40 miles apart. Soils at the Brazos Bottom are characterized by a higher water-holding capacity than those at the Stiles Farm. In addition, the Brazos Bottom has greater average annual and average growing season rainfall (39 and 23 inches compared to 36 and 19 inches).

Management practices varied within the pooled data set are maturity class, planting population, applied nitrogen rate, and planting date. All of the management practices were not varied within an individual study; however, when pooled, a diverse range of practices is obtained. Three maturity classes, medium, medium-late, and late season, are included in the data set. Planting populations range from 11,008 to 36,027 seeds/acre. Applied nitrogen varied from 40 to 267 lbs/acre. Planting dates ranged from February 8th to April 24th (Julian dates 39 to 114). The pooled data set contains 1,011 observations of which 583 are from the Brazos River Bottom location.

A more thorough discussion of the individual plot studies and the pooled data set can be found in Mjelde *et al.* One difference exists, however, between the data set discussed in Mjelde *et al.* and the data set used in this study. All observations in which no nitrogen was applied are deleted from Mjelde *et al.* data set to obtain the data set used here. This deletion is made so the logarithm of applied nitrogen could be taken for some of the heteroscedastic error models and for estimating the translog production function.

A priori it is reasonable to assume that heteroscedasticity exists in the pooled corn data set. Two locations are included in the data with slightly different climates and soil types. Further, each study measured the impact of only one or two management practices on yield. Finally, different researchers were involved in each of the plot studies. The pooled data set is, therefore, ideal for a large sample empirical study on the effects of heteroscedasticity.

Estimation of Production Functions

For each heteroscedastic error models, three different production functions are estimated. These three forms are modifications of the quadratic, translog, and square root functions, because interactions between some of the independent variables are excluded.

The quadratic production function to be estimated is:

$$
\begin{aligned}
Y = \beta_1 &+ \beta_2\, Nit + \beta_3 Nit^2 \\
&+ \beta_4 Pop + \beta_5 Pop_2 + \beta_6 Pd \\
&+ \beta_7 Pd^2 + \beta_8 Loc \\
&+ \beta_9 LocPop + \beta_{10} LocNit \\
&+ \beta_{11} MC1 + \beta_{12} MC2 \\
&+ \beta_{13} LocPd + \beta_{14} PopNit \\
&+ \beta_{15} PopPd + \beta_{16} NitPd
\end{aligned} \tag{10}
$$

where Y is corn yield (bu/ac), Nit is applied nitrogen (lbs/ac), *Pop* is planting population (thousand seeds/ac), *Pd* is planting date (Julian date), *Loc* is a 0-1 dummy for location (0 for Stiles Farm and 1 for Brazos Bottom), *MC1* is a 0-1 dummy for medium maturity classes, *MC2* is a 0-1 dummy for medium-late maturity classes, and β_i are parameters to be estimated[1]. To avoid a singular matrix, the late season maturity classification is the reference category.

Both the square root and translog production function are the same form as the quadratic form with appropriate redefinition of the variables. For the translog production function, *Y, Nit, Pop*, and *Pd* are redefined as the natural logarithm of yield, applied nitrogen, planting population, and planting date. Whereas, for the square root production function, these variables are redefined as the square root of applied nitrogen, planting population, and planting date. Yield is not redefined for the square root form. The dummy variables remain the same for all three functional forms.

Interactions between maturity class and the other management practices are omitted in (14). These interactions are omitted because of difficulties in defining the maturity classes (Mjelde *et al.*). Because of the absence of interactions, maturity class shifts yields but has no affect on the production- and profit-maximizing levels of applied nitrogen, planting population, or planting date.

As noted earlier, the exogenous variables in $h(x_t, \gamma)$ are precisely the same as those in $f(x_t, \beta)$. Such a general formulation subsumes any formulation where only a subset of the x_t's are used. Further, this formulation allows for differences between locations, because of the location variables. Finally, analyses by site indicated that heteroscedasticity exists in the data set within a site at p-values less that .05. As such, it is appropriate to model $h(x_t, \gamma)$ in the manner chosen.

Empirical Results

Problems occurred when estimating the translog and square root production function using error model number five. A negative variance was obtained for several of the observations when using the estimated error model five for these two functional forms. Recall, that the weight associated with model five is one over the square root of the estimated variance (see (5)). Because the square root of a negative number is complex, real-valued the translog and square root production functions could not be estimated using error model five. A negative variance occurred because no restrictions are placed on the possible values for predicted variances. Some error structures may, therefore, not be appropriate for certain situations.

Given the three alternative specifications of $f(x_t, \beta)$ and six alternative heteroscedasticity structures, full presentation of empirical results would be cumbersome. Consequently, only a summary of the results is presented. Additional results not presented here are given in Mjelde, Griffin, and Capps.

Production Function - Estimated Models

Summary statistics are given in table 2, for the various GLS estimated production functions. Each of the error models detected heteroscedasticity as given by the F-statistics associated with the variance model estimation. The large sample size may contribute to the detection of heteroscedasticity by all the error models.

The adjusted R^2 values for all the error models are low, suggesting weather or other omitted

Table 2. Summary of Production Function and Error Model Estimates

	Error Model[a]						
Factor	No. Het. Cor.[b]	One	Two	Three	Four	Five	Six
		Quadratic Production Function					
$f(x_i, \beta)$							
Adj. Corr.[c]	0.51	0.49	0.50	0 50	0 49	0.50	0.50
$\hat{\sigma}$	621.8	3.53	1.02	1.54	3.53	1.11	1.54
F-Test	69.9	66.2	68.3	67.5	65.3	67 4	67.5
5%[d]	13.	12.	11.	11.	11.	12.	11.
$h(x_i, \gamma)$							
$\bar{R}^2$		0.05	0.08	0.10	0.05	0.08	0.10
$\hat{\sigma}$		5 42	7.x10[5]	201.23	5.44	7.x10[5]	201 97
F-Test		4.9	7 0	8 6	4.6	7.0	8.4
		Translog Production Function					
$f(x_i, \beta)$							
Adj. Corr.[c]	0.52	0.45	0 49	0.48	0.45	NA[e]	0.48
$\hat{\sigma}$	0.07	3.30	.97	1.62	3.29	NA[e]	1.60
F-Test	74.0	56.0	65.5	62.6	55.0	NA[e]	62.6
5%[d]	13.	10.	10	11	9	NA[e]	10
$h(x_i, \gamma)$							
$\bar{R}^2$		0 16	0.27	0.29	0.15	0.27	0.29
$\hat{\sigma}$		4 40	0.012	0.02	4.41	0.01	0.02
F-Test		13.5	26 2	28.4	13.2	26.4	28 1
		Square Root Production Function					
$f(x_i, \beta)$							
Adj. Corr[c]	0.51	0.49	0 50	0 50	0 49	NA[e]	0 50
$\hat{\sigma}$	623 12	3.35	1.05	1 30	3.34	NA[e]	1.49
F-Test	69.6	65.3	67.9	67.8	64.7	NA[e]	67.4
5%[d]	15.	12.	13.	13.	11.	NA[e]	13.
$h(x_i, \gamma)$							
$\bar{R}^2$		0.06	0.08	0.10	0.06	0.08	0.10
$\hat{\sigma}$		4.66	7x10[5]	201.93	4.67	7x10[5]	202.24
F-Test		5.3	6.8	8.6	5.2	6.7	8.5

a See Table 1 for a definition of the heteroscedastic error models

b Estimated production function without correcting for heteroscedasticity.

c Adjusted squared correlation between the predicted and observed yields.

d Number of parameters, excluding the intercept, that are significant at the level of .05

e Unable to estimate this functional form, because a negative estimated variance occurred for some of the observations.

variables play a larger role in corn yield variance than the monitored input practices. For a given form of $f(x_i, \beta)$ the $\bar{R}^2$ for the error models are similar. The major difference between the error model estimates is in the estimated standard error of the regression $\hat{\sigma}$. These estimates vary dramatically among the error models. Examination of the error model standard goodness-of-fit tests, such as adjusted R^2, reveals no one error model clearly dominates the others.

Although the statistical summary of the models, adjusted R^2 values, estimated standard errors of regression, and F-tests are similar between the models (table 2), the estimated coefficients vary among the models. These differences are examined in the next two subsections.

Production Maximization

Production-maximizing level of inputs and associated yield for each production functional form and error model are listed in table 3 for Brazos Bottom and table 4 for Stiles Farm. For all models except the quadratic production function at the Stiles location, the corrected models maximize production at lower levels of applied nitrogen than the uncorrected models. This result is especially apparent for the translog form. The uncorrected translog form maximizes, for example, production

Table 3. Production-Maximizing Levels of Inputs and Associated Yields for a Late Season Hybrid, Brazos River Bottom Functions

Error Model [a]	Nit (lbs/ac)	Pop (thousand seeds/ac)	Pd (Julian Date)	YIELD (bu/ac)	Profit ($/ac)[b]
		Quadratic Production Function			
No Het. Cor [c]	186.6	27.452	51	147.6	306.90
One	164.4	26 986	65	149 8	317.30
Two	177.0	26 767	55	146.6	307.13
Three	172.9	26.589	59	147 5	310.21
Four	167.1	26.488	64	149.3	316 01
Five	179.3	26.856	53	147.0	307 41
Six	177.1	26 357	57	147.2	308 86
		Translog Production Function			
No Het. Cor [c]	328.7	21.581	52	144.7	296.60
One	143 1	24.685	63	146.6	315.44
Two	129.7	23 616	64	142 9	310.13
Three	150 9	24.228	60	145.4	311.44
Four	145 2	24.507	63	146.8	316.02
Five[d]					
Six	153 7	24.099	60	145.8	312.04
		Square Root Production Function			
No Het Cor.[c]	170.1	25.480	57	145.8	307.52
One	151.0	26.013	65	148.8	318.50
Two	166.3	25.509	57	145.2	306.86
Three	162.3	25.485	60	146.1	309 88
Four	152.0	25.837	65	149 0	318 79
Five[d]					
Six	164.2	25.374	59	146.1	309.51

a See Table 1 for a definition of the heteroscedastic error models.

b Profit (net of only applied nitrogen and seeding density) associated with the production maximizing level of inputs. Prices used are· corn price of $2 50/bushel, nitrogen price of $0 20/pound, and a seed price of $0.90/thousand seeds.

c Estimated production function without correcting for heteroscedasticity.

d Unable to estimate the production function using this error structure, because a negative estimated variance occurred for some of the observations.

at 2,583 lbs/A. of applied nitrogen at Stiles Farm. Clearly, this result is erroneous. For all functional forms, differences between the uncorrected model and corrected models for the remaining two management practices are not as pronounced as for nitrogen, but they do exist. The largest differences in planting population are for the translog form. Differences in planting dates also are evident. For the uncorrected forms, it is desirable to plant earlier (up to 14 days) than for the corrected functions for the Brazos Bottom location. Planting date differences are less pronounced for the Stiles location. For the uncorrected models, generally, it is desirable to plant later than for the corrected production functions for the Stiles location. These findings suggest that, if the objective is yield maximization not correcting for heteroscedasticity when it is present may lead to different input usage recommendations.

Differences between the corrected models also exist, but they are not as pronounced as between the uncorrected and corrected models. Further, differences exist among functional forms. Applied nitrogen ranges, for example, from 131-152 lbs/A. for the quadratic form depending on the error model compared to a range of 119-146 lbs/A. for the square root form for the Stiles location. For production maximization, the translog functional form has a smaller range of levels for planting date. No additional generalizable results exist for the remaining inputs between the two locations.

As noted earlier, the Brazos Bottom area has a higher average rainfall and better soil. These differences make the Brazos Bottom better suited to corn growth. As a consequence, more applied nitrogen and a higher planting population are recommended for the Brazos Bottom area (Cothren; Mjelde *et al.*). Given the close proximity of the two areas, planting dates might be expected to be similar. Planting dates between the two locations do differ for some of the error models under the quadratic form. These differences arise because the

Table 4. Production-Maximizing Levels of Inputs and Associated Yields for a Late Season Hybrid, Stiles Farm

Error Model [a]	Nit (lbs/ac)	Pop (thousand seeds/ac)	Pd (Julian Date)	YIELD (bu/ac)	Profit ($/ac)[b]
		Quadratic Production Function			
No Het. Cor.[c]	131 9	11 898	78	117 8	257.42
One	131.2	13 643	67	117.1	254 26
Two	143 8	13 355	71	116 3	249 94
Three	136 6	13.753	70	115.8	249.69
Four	131 6	15.313	68	113 0	242 34
Five	151 8	14.073	72	114.1	242 23
Six	139 5	14.945	70	112.4	239 56
		Translog Production Function			
No Het. Cor.[c]	2582 6	8.383	56	174 4	-88.14
One	127 5	14.433	61	115 2	249 41
Two	145.1	14 784	62	113 6	241 66
Three	137.5	14 402	62	116 5	250.68
Four	128 1	14 842	61	113 8	245 48
Five[d]					
Six	139 4	14.628	62	115.8	248 36
				Square Root Production Function	
No Het Cor[c]	151 6	14.093	66	117 0	249.60
One	120 0	15.150	63	116.0	252.26
Two	145 9	14 665	66	115 4	246 10
Three	138 3	14.937	64	114.9	246.18
Four	119 4	15.791	64	114 2	247 50
Five[d]					
Six	135 9	15.442	65	113 3	242.19

a See Table 1 for a definition of the heteroscedastic error models.

b Profits (net of only applied nitrogen and seed costs) associated with the production-maximizing level of inputs. Prices used are: corn price of $2.50/bushel, nitrogen price of $0.20/pound, and a seed price of $0.90/thousand seeds

c Estimated production function without correcting for heteroscedasticity.

d Unable to estimate the production function using this error structure, because a negative estimated variance occurred for some of the observations.

production-maximizing planting dates under the quadratic form for the Stiles Farm are generally later than for the other forms and location.

Profit Maximization

Profit-maximizing input levels are calculated for the following cost/price scenario. Corn price is assumed fixed at $2.50/bu, nitrogen costs $0.20/lb, and seeds cost $0.90/thousand seeds. No cost other than yield loss is associated with planting date. Profit-maximizing levels of inputs, associated profits, and yields are presented in tables 5 and 6 for Brazos Bottom and Stiles Farm. The profits reported are net of only applied nitrogen and seeding density costs.

It was noted earlier that, in all cases except the quadratic production function at Stiles, the uncorrected models maximized production at higher nitrogen levels than the corrected models. This situation generally does not hold for the profit-maximizing input levels given in tables 5 and 6. Most notably, the uncorrected translog production function has a more realistic profit-maximizing nitrogen level than what was calculated for production maximization. Noteworthy differences between profit-maximizing input levels exist among the corrected models.

The quadratic models for the Brazos Bottom, generally, have higher profit- maximizing levels of nitrogen and planting population than given by either the translog or square root models. Calculated profit-maximizing planting dates have a greater range for the Brazos Bottom than for the Stiles Farm. The profit-maximizing planting date for Stiles is later for the quadratic forms than for either the translog or square root functional form. Corn yields associated with the various models and profit-maximizing levels are similar within a given location.

Table 5. Profit-Maximizing Levels of Inputs, Associated Profit and Yield for a Late Season Hybrid, Brazos River Bottom[a]

Error Model[b]	Nit (lbs/ac)	Pop (thousand seeds/ac)	Pd (Julian Date)	PROFIT ($/ac)	YIELD (bu/ac)[c]
		Quadratic Production Function			
No Het Cor.[d]	138 6	26.169	62	312.28	145.4
One	133 2	26 587	69	320 60	148.5
Two	143 1	26.153	60	310.79	145.2
Three	140.0	26.046	64	313 74	146.1
Four	134 4	26 167	69	319.42	147 9
Five	146 9	26 414	57	310.85	145.6
Six	142 8	25.884	62	312.51	145 7
		Translog Production Function			
No Het Cor.[d]	123 6	22.712	63	304.94	140.0
One	116 7	24.532	66	318 40	145 5
Two	94.6	23.277	70	314 44	141.7
Three	122.9	24.123	63	314.63	144.4
Four	117 8	24 422	66	319.05	145.8
Five[f]					
Six	124 7	24.052	62	315.35	144 8
		Square Root Production Function			
No Het. Cor[e]	118 8	25 189	65	313.19	143 9
One	123 0	25.756	69	321 56	147.7
Two	129 7	25.346	61	310.86	143.8
Three	125 7	25.298	64	313.85	144 7
Four	123.0	25.640	69	321 94	147.8
Five[f]					
Six	127.5	25.226	63	313 54	144.7

a Prices used in maximizing profits for a late season hybrid are: corn price of $2.50/bushel, nitrogen price of $0.20/pound, and a seed price of $0.90/thousand seeds

b See Table 1 for a definition of the heteroscedastic models.

c Yield associated with the profit-maximizing level of inputs

e Estimated production function without correcting for heteroscedasticity.

f Unable to estimate the production function using this error structure, because a negative estimated variance occurred for some of the observations.

Between the error models within a functional form grouping, notable differences exist for the profit-maximizing levels of inputs. Applied nitrogen levels for Brazos Bottom, for example, ranges from 95 lbs/A. to 125 lbs/A. with the translog functional form. This difference of 30 lbs/A. is between one-third and one-quarter of the profit-maximizing input level (depending on the base used). On the other extreme, differences in planting dates for Stiles vary only 4 to 6 days depending on functional form.

Similar to production maximization, it is difficult to discern general conclusions concerning profit-maximizing input levels. When comparing the production- and profit-maximizing planting date ranges for a given location, few differences exist. This lack of changing planting dates is most likely because no economic cost other than yield loss is associated with planting date. Few generalities are evident concerning the other two inputs when comparing production- and profit-maximizing ranges of inputs.

The above discussion on production- and profit-maximization levels of inputs and resulting yields and profits provides a comparison between the functions at essentially a single point. To provide a more comprehensive comparison of the production functions, profit contours are used (as suggested by Debertin). Selected contours in nitrogen-planting date space (figures 1 and 2) are presented for the Brazos Bottom location. Additional contours including contours for the Stiles location can be found in Mjelde, Griffin, and Capps.

All contour plots were created using a similar procedure. The lowermost contour is $270/ac. The increment of each successive contour line is $5/ac until the profit-maximizing level of profit is obtained. Note that the maximum profit

Table 6 Profit-Maximizing Levels of Inputs, Associated Profit and Yield for a Late Season Hybrid, Stiles Farm[a]

Error Model[b]	Nit (lbs/ac)	Pop (thousand seeds/ac)	Pd (Julian Date)	PROFIT ($/ac)	YIELD (bu/ac)[c]
		Quadratic Production Function			
No Het. Cor [d]	83 9	10.615	89	262.79	115.7
One	100.0	13.244	72	257.56	115.8
Two	109.9	12 741	76	253.61	114 8
Three	103 8	13.210	75	253.21	114.3
Four	99.0	14 993	73	245 74	111 6
Five	119 4	13.630	76	245.67	112.7
Six	105.2	14.471	75	243 20	110 9
		Translog Production Function			
No Het. Cor [d]	281.96	10 047	84	319 74	154 1
One	101.2	14.477	64	252.30	114.2
Two	95 0	14 670	70	247.64	111 9
Three	107.9	14.467	65	253.80	115 4
Four	101.1	14 913	65	248 43	112 8
Five[e]					
Six	109.3	14 716	65	251.56	114.7
		Square Root Production Function			
No Het. Cor [d]	106.9	13 997	74	254 25	115 3
One	97.5	15 050	67	254.69	115 1
Two	115 1	14 637	69	249 21	114 2
Three	105 9	14.907	69	249.80	113 8
Four	96 3	15 716	68	249.95	113.3
Five[e]					
Six	105 3	15.410	69	245.47	112 2

a Prices used in maximizing profits for a late season hybrid are: corn price of $2 50/bushel, nitrogen price of $0 20/pound, and a seed price of $0.90/thousand seeds.

b See Table 1 for a definition of the heteroscedastic models

c Yield associated with the profit-maximizing level of inputs

d Estimated production function without correcting for heteroscedasticity.

e Unable to estimate the product function using this error structure, because a negative estimated variance occurred for some of the observations, thus, a square root could not be taken.

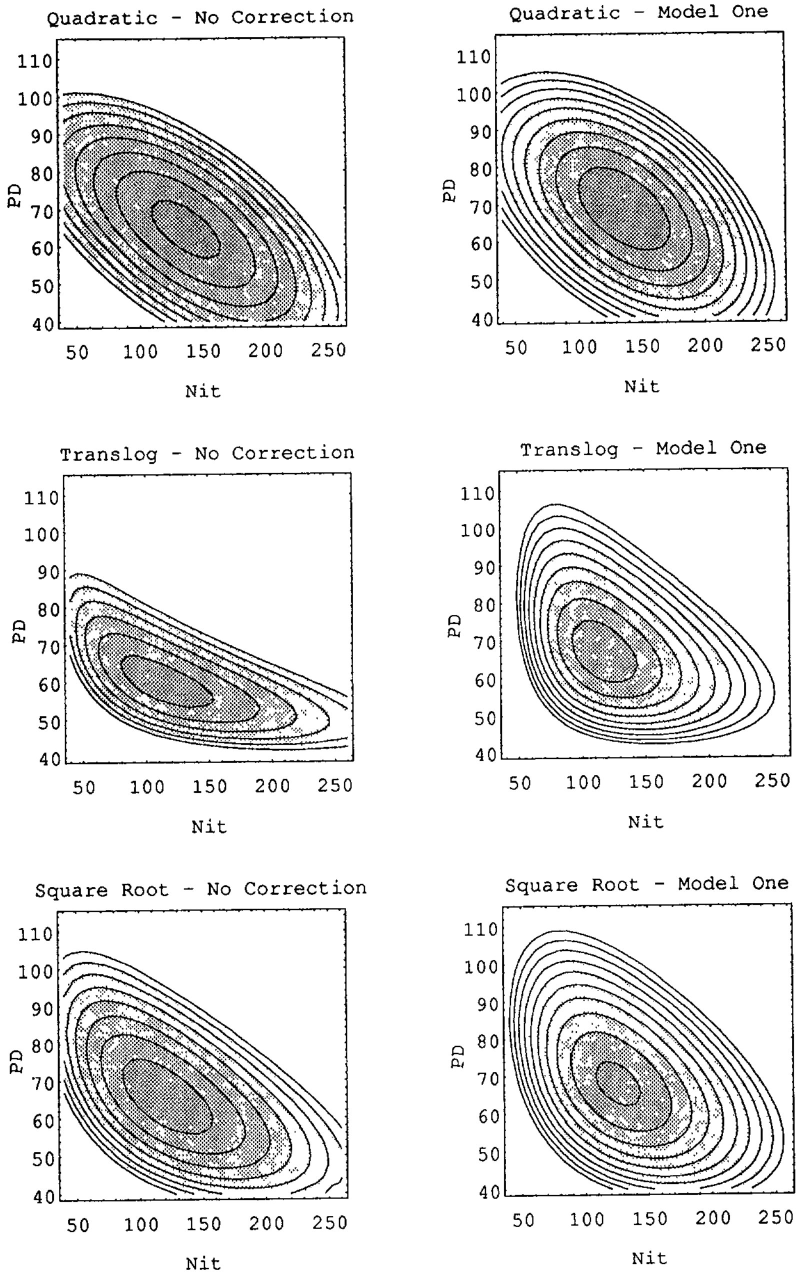
level varies by functional form and error model. Plotting range for the two inputs corresponds to the range of input levels within the data set. Because the plots are two-dimensional, a fixed level for the third input must be assumed. The input level for population is 25,000 seeds/ac which approximates the profit-maximizing level for all error models.

Only general observations and a few specific examples are discussed. Between the contours, the largest differences, generally, are between the functional forms. The next largest differences are between the uncorrected and corrected profit contours. Smaller relative differences are noted between the different error models. The translog production functional form has the largest differences of the three forms between the uncorrected and corrected profit contours. Contours associated with the translog and square root functional forms are, generally, less circular than for the quadratic form.

These general differences in profit contours have implications concerning input recommendations. The penalty associated with being further away from the profit-maximizing level of inputs differs between functions. Using different functions may result in different input recommendations especially when risk aspects and the stochastic nature of production are factored into the analysis. Differing recommendations would be a function of the steepness of the production function.

As noted, the profit contours provide information on the shape of the underlying functions. Consider, as an example of the steepness of the contours, error model one (figure 1). For the three functional forms, maximum profit is between $318 - $321/A. for the Brazos Bottom location. "Optimal" planting date is approximately Julian date 68 for all three forms, profit-maximizing levels of applied nitrogen are approximately 133, 116, and

Figure 1. Iso-profit Contours Contrasting Error Model One to the Production Function with No Correction for Heteroscedasticity for the Brazo



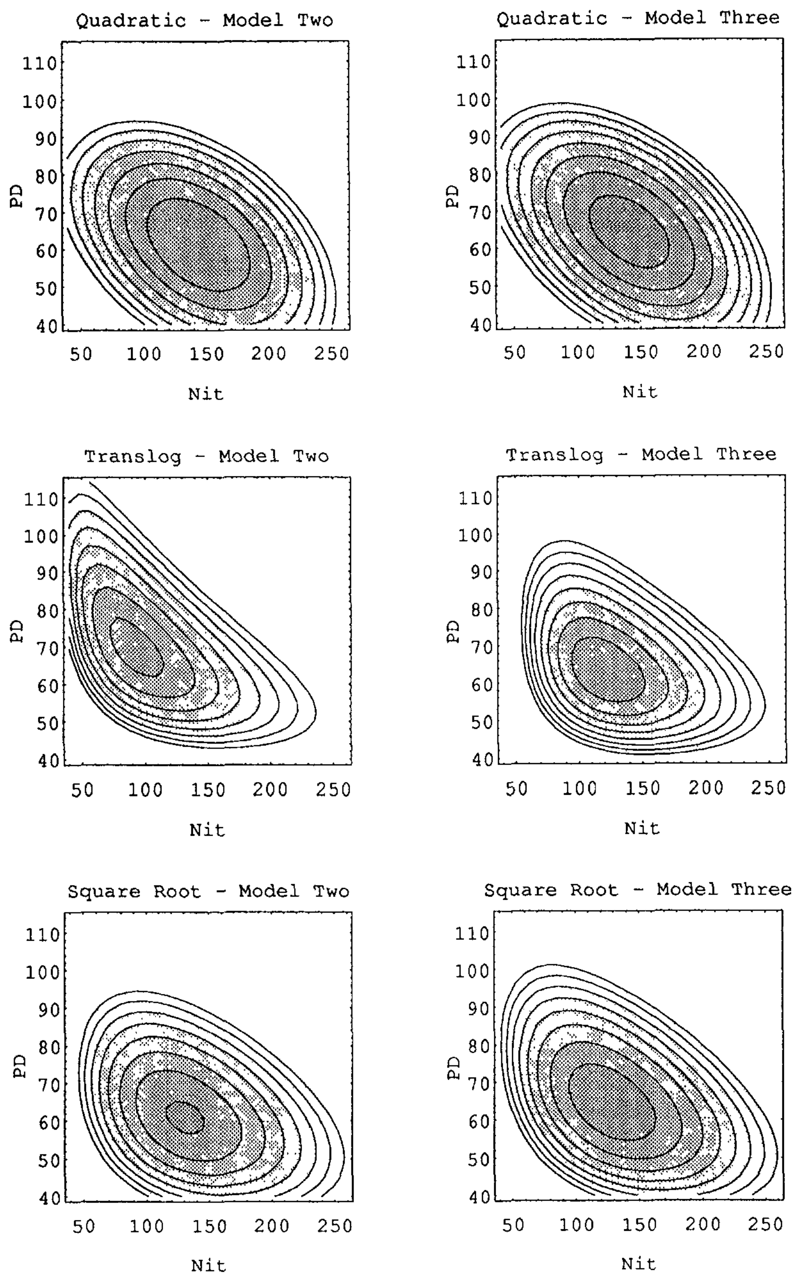
123 for the quadratic, translog, and square root production functions. Contours presented in figure 1 indicate that the profit surface is much steeper on the low input side for either the translog or square root function than it is for the quadratic. This result indicates that being further away from the profit-maximizing level of inputs decreases profits faster for either the square root or translog than for the quadratic.

Guidelines for Empirical Applications

In commenting on issues pertaining to the appropriate functional form, Hildreth observed:

> "it is particularly disconcerting that, in many instances in which several alternative assumptions [as to functional form] have been

Figure 2. Iso-profit Contours Contrasting Error Models Two and Three for the Brazos Bottom Location



investigated, alternative fitted equations have resulted which differ little in terms of conventional statistical criteria such as multiple correlation coefficients or F tests of the deviation, but differ much in their economic implication" (p. 64).

The results presented here extends Hildreth's observations concerning functional form to include problems of heteroscedasticity. Previous studies have examined the implications or sensitivity of choosing one functional form (e.g. Bay and Schoney; Griffin *et al.*). The conclusion of these previous studies is summarized by Griffin, Montgomery, and Rister:

> "given the possible differences in economic implications, it is often advisable to explore the sensitivity of calculated economic optima to the choice of functional form" (p. 224).

The results presented in this study support and extend their summary.

Based on the extension of previous findings provided by the current study, several recommendations for empirical studies are made. It is stressed these recommendations serve as a guideline and not as a cookbook. As with all empirical applications, prior knowledge and experience play an important role. Recommendations are:

(1) employ different production functional forms (e.g. quadratic, translog, and square root),

(2) use alternative tests for heteroscedasticity (e.g. Park-Glejser, Harvey, Breusch-Pagan-Godfrey),

(3) if heteroscedasticity exists, make appropriate corrections in the estimation procedure, and

(4) use both statistical and economic criteria in examining the consequences of different functional forms and heteroscedastic correction measures.

It should be noted that negative variances may arise in the estimation procedure and/or use of the production functions. This potential problem may limit which forms of heteroscedasticity can be examined.

As shown in the estimated production functions in this study, conventional statistical criteria may be similar among production functions and heteroscedasticity corrections, but optimal input recommendations may vary. In the functions used here, optimal input differences of up to one-third are noted. Further, to our knowledge, no statistical tests exist which jointly test for both functional form and form of heteroscedasticity. Nonnested tests of hypothesis may be a possibility, but this area remains for future research. The "best" procedure is to simply report ranges of optimal input levels or the economic implications associated with the different functional forms and forms of heteroscedasticity. In studies where it is impractical to provide the sensitivity results, at a minimum, potential biases should be noted along with the rationale used to select the chosen function.

Discussion and Implications

Several implications can be drawn from the results of this empirical study of heteroscedasticity. First, this study confirms previous studies concerning the need to correct for heteroscedasticity. Production- and profit-maximizing levels of inputs and the shape of the profit equation obtained from the uncorrected models differed from those obtained from the corrected models. In some cases, the levels obtained from the uncorrected models were nonsensical. This result is especially apparent for the translog production function. Given the need to correct for heteroscedasticity, the means by which the corrections are obtained should not be undertaken lightly. One should not merely correct for heteroscedasticity without exploring alternative specifications and the implications of the selected specification. Differences of up to one-quarter to one-third the amount of production- or profit-maximizing levels of inputs are calculated when using the different error correction models.

Results both support and contradict Judge *et al.* (1985)'s hypothesis that the choice of which heteroscedasticity correction to employ may be unimportant. If the model is used for prediction purposes, that is to predict yields or profits, all error correction models gave approximately the same levels. These results support Judge *et al. (1985)*'s hypothesis. Results concerning the production/profit maximizing levels of inputs contradict Judge *et al.*

(1985)'s hypothesis when using the models in a recommendation rather than a predictive framework. Additional studies need to be conducted concerning these two uses for models and correcting for heteroscedasticity.

The findings of this analysis are restricted to the data set used but should generate potential discussion concerning heteroscedasticity corrections. A natural extension of this study is to examine various known forms of heteroscedastic error terms in a Monte Carlo study. Further, other functional forms including flexible forms, such as the Fourier form, warrant examination. Large sample *versus* small sample properties should be examined.

References

Amemiya, T. "Regression Analysis When the Variance of the Dependent Variable is Proportional to the Square of its Expectation." *J. Amer. Stat. Association* 68(1973):928-34.

Amemiya, T. "A Note on a Heteroscedastic Model." *J. of Econometrics* 6(1977):365-70.

Battese, G.E. and B.P. Bonyhady. "Estimation of Household Expenditure Functions: An Application of a Class of Heteroscedastic Regression Models." *The Economic Record* 57(1981):80-5.

Bay, T.F. and R.A. Schoney. "Data Analysis with Computer Graphics: Production Functions." *Amer. J. Agr. Econ.* 64(1982):289-97.

Breusch, T.S. and A.R. Pagan. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrics* 47(1979):1287-94.

Buccola, S.T. and B.A. McCarl. "Small-Sample Evaluation of Mean-Variance Production Function Estimators." *Amer. J. of Ag. Econ.* 68(1986):732-8.

Carroll, R.J. and D. Ruppert. "Transformation and Weighting in Regression." New York. Chapman and Hall. 1988. pp. 249.

Cothen, J.T. Personnel Communications. Professor, Soil and Crops Department. Texas A&M University. June 1989.

Debertin, D.L. "An Animated Instructional Module for Teaching Production Economics with 3-D Graphics." *Amer. J. Ag. Econ.* 75(1993):485-491.

Glejser, H. "A New Test for Heteroscedasticity." *J. Amer. Stat. Association* 64(1969):316-23.

Griffin, R.C., J.M. Montgomery, and M.E. Rister. "Selecting Functional Form in Production Function Analysis." *W. J. Ag. Econ.* 12(1987):216-227.

Griffin, R.C., M.E. Rister, J.M. Montgomery, and F.T. Turner. "Scheduling Inputs with Production Functions: Optimal Nitrogen Programs for Rice." *S. J. Agr. Econ.* 17(1985):159-68.

Harvey, A.C. "Estimating Regression Models with Multiplicative Heteroscedasticity." *Econometrica* 44(1976):461-5.

Hildreth, C. and J.P. Houck. "Some Estimators for a Linear Model with Random Coefficients." *J. Amer. Stat. Association* 63(1968):584-95.

Hildreth, C.G. "Discrete Models with Qualitative Descriptions." *Methodological Procedures in the Economic Analysis of Fertilizer Use Data*, ed. E.L. Baum, E. O. Heady, and J. Blackmore, pp. 62-75. Ames:Iowa State College Press, 1955.

Judge, G.C., R.C. Hill, W.E. Griffiths, H. Lütkepohl, and T.C. Lee. *The Theory and Practice of Econometrics.* First Edition. John Wiley and Sons, 1980.

Judge, G.C., R.C. Hill, W.E. Griffiths, H. Lütkepohl, and T.C. Lee. *Introduction to the Theory and Practice of Econometrics.* John Wiley and Sons, 1982.

Judge, G.C., W.E. Griffiths, R.C. Hill, H. Lütkepohl, and T.C. Lee. *The Theory and Practice of Econometrics.* Second Edition, John Wiley and Sons, 1985.

Just, R.E. and R.D. Pope. "Stochastic Specification of Production Functions and Economic Implications." *J. Econometrics* 7(1978):67-86.

Just, R.E. and R.D. Pope. "Production Function Estimation and Related Risk Considerations." *Amer. J. Ag. Econ.* 61(1979):276-84.

Maddala, G.S. *Econometrics.* New York: McGraw-Hill Book Company. 1977.

McCarl, B.A. and R.B. Rettig. "Influence of Hatchery Smolt Releases on Adult Salmon Production and Its Variability." *Canadian J. Fisheries and Aquatic Science* 40(1983):1880-6.

Mjelde, J.W., J.T. Cothern, M.E. Rister, E.M. Hons, C.G. Coffman, C.R. Shumway, and R.G. Lemon. "Integrating Data from Various Field Experiments: The Case of Corn in Texas." *J. Prod. Ag.* 4(1991):139-147.

Mjelde, J.W., R.C. Griffin, and O. Capps, Jr. "Empirical Considerations of Heteroscedasticity." Dept. Information Report DIR 93-1, SP-3, Dept. of Ag. Econ. Texas A&M University. 1993.

Paris, S.J. "A Note on Heteroscedasticity Errors in Regression Analysis." *Review of the International Statistical Institute* 21(1953):28-9.

Park, R.E. "Estimation with Heteroscedastic Error Terms." *Econometrica* 34(1966):888.

Yang, S.R., W.W. Koo, and W.W. Wilson. "Heteroscedasticity in Crop Yield Models." *West. J. Ag. Econ.* 17(1992):103-109.

Endnote

1. Formulations which included both growing season rainfall and temperature and only rainfall were considered. A high degree of multicollinearity existed because the climatic variables vary only across years (only five years) and locations. Location and rainfall are associated in the multicollinearity. Because of this problem, it is felt that the formulation presented provides "better" estimates of the individual parameters to be used in the optimization.