ESTIMATION OF A PRODUCTION FRONTIER MODEL: WITH APPLICATION TO THE PASTORAL ZONE OF EASTERN AUSTRALIA

GEORGE E. BATTESE AND GREG. S. CORRA*

University of New England

and

Bureau of Agricultural Economics

This paper considers a statistical model for a production frontier that is consistent with the traditional (nonstochastic) definition of a production function given in microeconomic theory. Limiting cases of the model are the familiar average production function and an envelope production function. Maximum-likelihood estimators for the parameters of the model are defined. The three related models are applied in the estimation of a production frontier for the Pastoral Zone of Eastern Australia with use of data from the Australian Grazing Industry Survey.

Introduction

The concept of a 'production function' is basic to the development of the theory of the firm in microeconomic theory. In the classical non-stochastic theory of the firm a production function is defined as '. . . a schedule showing the maximum amount of output that can be produced from a specified set of inputs, given the existing technology' (Ferguson [7], p. 110). Several statistical models, that are consistent with this definition of a production function, have been discussed in recent literature (e.g. Aigner and Chu [2], Aigner, Amemiya and Poirier [1], and Aigner, Knox Lovell and Schmidt [3]).

Aigner and Chu [2] considered a model in which output observations never exceeded the corresponding value on the production function. Deviations of output observations from the production function were assumed due to technical inefficiency of a firm or detrimental occurrences experienced during the production process. Timmer [10] suggested that, since the presence of outliers is likely to significantly change the estimated production function, a predetermined proportion of output observations should be permitted to lie above the estimated production function.

Aigner, Amemiya and Poirier [1] and Aigner, Knox Lovell and Schmidt [3] have suggested statistical models for production functions

---

that permit output observations to lie above the production function. The random errors in the Aigner, Amemiya and Poirier [1] model are a mixture of positive and negative truncated normal random variables. The Aigner, Knox Lovell and Schimdt [3] model is discussed below, although a slight reparameterization is suggested for estimation purposes.

The Production Frontier Model

The statistical model for the production function that we consider is defined by

$$Y_t = x_t\beta + E_t, \quad t = 1, 2, \ldots, T \tag{1}$$

such that

$$E_t = U_t + V_t, \quad t = 1, 2, \ldots, T \tag{2}$$

where

$Y_t, t = 1, 2, \ldots, T$, are observable random variables that are output values;

$x_t, t = 1, 2, \ldots, T$, are $(1 \times k)$ vectors of non-negative constants that are input values such that the first element of x_t is one and the matrix $(\sum_{t=1}^{T} x'_t x_t)$ is non-singular;

β is a $(k \times 1)$ vector of unknown constants; and

U_t and $V_t, t = 1, 2, \ldots, T$, are unobservable random errors that are independently distributed.

The random errors, $U_t, t = 1, 2, \ldots, T$, in (2) are assumed to be negative and arise by truncation of the normal distribution with mean zero and positive variance σ_U^2, whereas the random errors, $V_t, t = 1, 2, \ldots, T$, are assumed to have normal distribution with mean zero and positive variance σ_V^2. The density function for U_t is defined by

$$f_{U_t}(u_t) = 2(2\pi\sigma_U^2)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}u_t^2/\sigma_U^2), \quad \text{if } u_t < 0$$
$$= 0, \quad \text{if } u_t \geq 0. \tag{3}$$

The negative error, U_t, in the model is interpreted with reference to a firm's technical inefficiency of production, while the symmetric error, V_t, which can be considered a 'measurement error', is associated with uncontrollable factors related to the production process. The presence of the error, V_t, implies that the output value, $x_t\beta + U_t$, is not observable.

If the random errors, V_t, are absent from the error model (2), then the model obtained is that considered by Aigner and Chu [2], which we call the 'full-frontier model'. In this case the range of the dependent random variable depends on the β-parameters, i.e., $y_t \leqslant x_t\beta, t = 1, 2, \ldots, T$.

If the random errors, U_t, are absent from the model (2), then the model obtained is called the 'average-frontier model', the one most often estimated in econometric studies. For the production function (1) having errors, $E_t = V_t, t = 1, 2, \ldots, T$, the expectation of output observations for a given vector of input values, x_t, is $x_t\beta$, and under the stated assumptions the output observations exceed the expectation with probability 0·5.

The production function (2), in which the variances σ_U^2 and σ_V^2 are positive, is referred to in this paper as the 'pseudo-frontier model'. The average- and full-frontier models are limiting cases of the pseudo-frontier model.

The density function for the observable random variable, Y_t, is given by

$$f_{Y_t}(y_t) = f_{E_t}(y_t - x_t\beta) \tag{4}$$

where $f_{E_t}(\,.\,)$ denotes the density function for E_t. The density function for E_t is obtained from the joint density function for U_t and V_t by the change-of-variable technique that involves a straightforward but somewhat tedious exercise in integration. That is, by the transformation $E_t = U_t + V_t$ and (say) $W_t = V_t$, where $E_t < W_t$, the joint density function for E_t and W_t evaluated at e_t and w_t is the product of the marginal density functions for U_t and V_t evaluated at $e_t - w_t$ and w_t, respectively. Thus the density function for E_t is defined by

$$f_{E_t}(e_t) = \int_{e_t}^{\infty} f_{U_t}(e_t - w_t) f_{V_t}(w_t) dw_t \tag{5}$$

where $f_{U_t}(\,.\,)$ is defined by (3) and $f_{V_t}(\,.\,)$ is the density function for the normal distribution with mean zero and variance σ_V^2. The solution to the integral equation (5) is given by (see Corra [5, pp. 13–15])

$$f_{E_t}(e_t) = 2[1 - \Phi(e_t\sigma_U/\sigma_V\sigma)](2\pi\sigma^2)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}e_t^2/\sigma^2) \tag{6}$$

where $\sigma^2 = \sigma_U^2 + \sigma_V^2$ and $\Phi(\,.\,)$ denotes the distribution function for the standard normal random variable,

i.e., $\qquad \Phi(z) = \int_{-\infty}^{z} (2\pi)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}t^2)dt.$

Equations (4) and (6) imply that the density function for Y_t can be expressed by

$$f_{Y_t}(y_t) = 2[1 - \Phi(z_t)](2\pi\sigma^2)^{-\frac{1}{2}}\exp[-\tfrac{1}{2}(y_t - x_t\beta)^2/\sigma^2] \tag{7}$$

where

$$z_t = [(y_t - x_t\beta)/\sigma][\gamma/(1 - \gamma)]^{\frac{1}{2}} \tag{8}$$

and

$$\gamma = \sigma_U^2/\sigma^2. \tag{9}$$

The expression for the density function for Y_t, given by (7), is not identical to that given by Aigner, Knox Lovell and Schmidt [3, equation (8)] in that the latter authors express the density function in terms of $\lambda \equiv \sigma_U/\sigma_V$ instead of γ. The parameter γ is bounded between zero and one, whereas the parameter λ can be any positive real number. We suggest that estimation of the model in terms of the formulation (7)-(9) is computationally preferable.

It is noted that if the variance σ_U^2 approaches zero, then the variable z_t approaches zero and so the factor $2[1 - \Phi(z_t)]$ in (7) approaches one. In this case the density function for Y_t approaches that of a normal distribution with mean $x_t\beta$ and variance $\sigma^2 \equiv \sigma^2_V$. This is the density function for the 'average-frontier model'. If, however, the variance σ^2_V approaches zero, then the variable z_t approaches minus infinity and the factor $2[1 - \Phi(z_t)]$ in (7) approaches two. In this case the density function for Y_t approaches that of the truncated normal distribution that is obtained from (3) by substituting $y_t - x_t\beta$ for u_t.

It should be noted that the parameter σ_U^2 is not the variance of the negative error, U_t. It can be shown that $E(U_t) = -(2\sigma_U^2/\pi)^{\frac{1}{2}}$ and Var $(U_t) = \sigma_U^2 (\pi - 2)/\pi$. It therefore follows that the expectation and variance of Y_t in terms of our parameterization of the model are

$$E(Y_t) = x_t\beta - (2\gamma\sigma^2/\pi)^{\frac{1}{2}}$$

and

$$\text{Var}(Y_t) = \sigma^2(\pi - 2\gamma)/\pi.$$

Given the sample observations, $y_1, y_2, \ldots, y_T$, on the random variables, $Y_1, Y_2, \ldots, Y_T$, the logarithm of the likelihood function is given by

$$L(\theta;y) \equiv -\tfrac{1}{2}T \ln (\tfrac{1}{2}\pi) - \tfrac{1}{2}T \ln \sigma^2 + \sum_{t=1}^{T} \ln[1 - \Phi(z_t)] -$$

$$\tfrac{1}{2} \sum_{t=1}^{T} (y_t - x_t\beta)^2/\sigma^2, \tag{10}$$

where $\theta \equiv (\beta', \sigma^2, \gamma)'$ denotes the $(k + 2)$ column vector of parameters of the model.

The maximum-likelihood estimates for the parameters of the model, denoted by $\hat{\theta}$, are defined by the solution to the $(k + 2)$ likelihood equations that are defined by equating the partial derivatives of the logarithm of the likelihood function to zero, i.e.,

$$\partial L(\hat{\theta};y)/\partial\theta = \underline{0} \tag{11}$$

where $\underline{0}$ denotes the $(k + 2)$ column vector of zeros.

The first-order partial derivatives of $L(\theta;y)$ with respect to β, σ^2 and γ can be shown to be

$$\frac{\partial L(\theta; y)}{\partial\beta} = \left[\frac{\gamma}{(1 - \gamma)\sigma^2}\right]^{\frac{1}{2}} \sum_{t=1}^{T} h(z_t)x_t' + \left[\frac{(1 - \gamma)}{\gamma\sigma^2}\right]^{\frac{1}{2}} \sum_{t=1}^{T} z_t x_t' \tag{12}$$

$$\frac{\partial L(\theta; y)}{\partial\sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^2} \sum_{t=1}^{T} h(z_t)z_t + \frac{(1 - \gamma)}{2\gamma\sigma^2} \sum_{t=1}^{T} z_t^2 \tag{13}$$

and

$$\frac{\partial L(\theta; y)}{\partial\gamma} = -\tfrac{1}{2}\left[\gamma(1 - \gamma)\right]^{-1} \sum_{t=1}^{T} h(z_t)z_t \tag{14}$$

where $h(z_t) \equiv \phi(z_t)/[1 - \Phi(z_t)]$

and $\phi(z_t) \equiv \dfrac{d}{dz_t} \Phi(z_t) = (2\pi)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}z_t^2)$.

It is clear from equations (12), (13) and (14) that the maximum-likelihood estimates for the parameters cannot be expressed in closed form and that they can only be approximated by numerical methods.

The well-known Newton-Raphson method is used to approximate the maximum-likelihood estimates of the parameters in the production frontier (1)–(2). Given an initial estimate, $\hat{\theta}_0$, for the vector θ, a new estimate, $\hat{\theta}_1$, is obtained by the Newton-Raphson technique according to the formula

$$\hat{\theta}_1 = \hat{\theta}_0 - \left[\frac{\partial^2 L(\hat{\theta}_0; y)}{\partial\theta\, \partial\theta'}\right]^{-1} \frac{\partial L(\hat{\theta}_0; y)}{\partial\theta} \tag{15}$$

where $\dfrac{\partial^2 L(\theta_0 ; y)}{\partial\theta\,\partial\theta'}$ denotes the square matrix of dimension $(k + 2)$, consisting of the second partial derivatives of the logarithm of the likelihood function, evaluated at the initial parameter estimates.

The initial estimate, $\hat\theta_0$, should be in a neighbourhood of the global maximum of the likelihood function for the Newton-Raphson estimator to have desirable statistical properties. For one to have reasonable assurance that approximate maximum-likelihood estimates are obtained, a wide search of the parameter space is required and the value of the likelihood function evaluated at different points. This is not an exceptionally difficult task for the production frontier model (1) because only three parameters involve any real problem of choosing reasonable initial estimates. These parameters are the intercept, denoted by β_o, and the variance parameters σ^2 and γ. The ordinary least-squares estimators for the other β-parameters are unbiased whereas the ordinary least squares estimator for the constant, β_o, unbiasedly estimates $\beta_o - (2\sigma_U^2/\pi)^{\frac{1}{2}}$ and the residual mean square from the regression unbiasedly estimates $\sigma^2 - (2\sigma_U^2/\pi)$. It is suggested that different initial estimates for γ, ranging from say $0\cdot1$ to $0\cdot9$, be used.

Given suitable regularity conditions (e.g., Amemiya [4] and Theil [9, p. 392]), it follows that the maximum-likelihood estimator, denoted by $\hat\theta_T$, is such that the random vector, $\sqrt{T}(\hat\theta_T - \theta)$, converges in distribution to a $(k + 2)$-variate normal random variable with zero mean vector and covariance matrix $\left\{\underset{T\to\infty}{\text{Lim}}\,E\left[-\dfrac{1}{T}\dfrac{\partial^2 L(\theta ; Y)}{\partial\theta\,\partial\theta'}\right]\right\}^{-1}$, where Y denotes the $(T\times1)$ vector of random variables $Y_1, Y_2, \ldots, Y_T$. The covariance matrix of the maximum-likelihood estimator is estimated by

$$\text{Cov}(\hat\theta_T) = \left[-\dfrac{\partial^2 L(\hat\theta_T ; Y)}{\partial\theta\,\partial\theta'}\right]^{-1}. \tag{16}$$

The second-order partial derivatives of $L(\theta ; y)$ are not given in this paper but they are presented by Corra [5].

It is noted that one of the basic regularity conditions to establish the above asymptotic result is that the expectations of the first-partial derivatives of the logarithm of the likelihood function are zero,

$$\text{i.e.,}\qquad E\left[\dfrac{\partial L(\theta ; Y)}{\partial\theta}\right] = \underline{0}. \tag{17}$$

This condition is satisfied for the pseudo-frontier model defined by (1) and (2). Proof that the result of (17) is true is obtained with use of the expectations

$$E(Z_t) = -(2/\pi)^{\frac{1}{2}}\gamma/(1 - \gamma)^{\frac{1}{2}}.$$
$$E\{\phi(Z_t)/[1 - \Phi(Z_t)]\} = [2(1 - \gamma)/\pi]^{\frac{1}{2}},$$

and

$$E\{\phi(Z_t)Z_t/[1 - \Phi(Z_t)]\} = 0,$$

where $Z_t = [(Y_t - x_t\beta)/\sigma][\gamma/(1 - \gamma)]^{\frac{1}{2}}$. These expectations are not proven here. Further, we have not been successful in obtaining the expectations of the second-order partial derivatives of the log-likelihood

function. Thus the inverse of the negative of the matrix of second-order partial derivatives, evaluated at maximum-likelihood estimates, is used [equation (16)] to estimate the covariance matrix of the maximum-likelihood estimators for the parameters of the production frontier. This estimated covariance matrix is, however, not necessarily positive definite.

· The condition of equation (17) is not satisfied for the 'full-frontier model' in which $\sigma^2_V = 0$. For this model the log-likelihood function is

$$L(\theta^*;y) = -\tfrac{1}{2}T \ln (\tfrac{1}{2}\pi) - \tfrac{1}{2}T \ln (\sigma_U{}^2) - \tfrac{1}{2} \sum_{t=1}^{T} (y_t - x_t\beta)^2/\sigma_U{}^2 \qquad (18)$$

where $y_t \leqslant x_t\beta$, $t = 1, 2, \ldots, T$, and
$\theta^* = (\beta', \sigma_U{}^2)'$.

From this expression it is clear that the maximum-likelihood estimate for β is obtained by minimizing the sum of squares, $\sum_{t=1}^{T} (y_t - x_t\beta)^2$, subject to the restrictions, $y_t \leqslant x_t\beta$, $t = 1, 2, \ldots, T$. This is clearly a quadratic programming problem.[1] The standard errors for the maximum-likelihood estimators for the β-parameters in this model are not obtained by use of (16).

Empirical Results

We apply the statistical model (1) in the estimation of a production frontier for sheep production in the Pastoral Zone of Eastern Australia which includes parts of the three States, South Australia, New South Wales, and Queensland. The model is defined by

$$Y_t = \beta_0 + x_{1t}\beta_1 + x_{2t}\beta_2 + x_{3t}\beta_3 + x_{4t}\beta_4 + x_{5t}\beta_5 + E_t,$$
$$t = 1, 2, \ldots, T. \qquad (19)$$

where the subscript t denotes the tth farm in the survey data consisting of T farms in the region;

Y_t denotes the logarithm of the value of sheep production;
x_{1t} denotes the logarithm of the cost of watering facilities;
x_{2t} denotes the logarithm of the cost of fencing;
x_{3t} denotes the logarithm of the cost of labour;
x_{4t} denotes the logarithm of the cost of machinery inputs;
x_{5t} denotes the logarithm of the costs associated with land;

and the random errors, E_t, $t = 1, 2, \ldots, T$, are assumed to be independent and identically distributed and have the same marginal distribution as the errors, E_t, of (2).

The values of the variables in the model (19) were calculated from data obtained by the Bureau of Agricultural Economics in the 1973-74 Australian Grazing Industry Survey. The definitions of the variables in terms of the variables of the BAE code for the survey are given in Corra [5]. The dependent variable is obtained by summing the net proceeds from wool sales and the net trading and inventory gain in livestock numbers during the financial year 1973-74. The value of the 'water variable', x_1, is calculated by summing the costs of improvements to the water supply, depreciation in the value of existing water supply

---

[1] This result is stated by Schmidt [8]. In addition, if the negative errors, U_t, in a full-frontier model have exponential distribution, then the maximum-likelihood estimates for the β-parameters are obtained by linear programming, namely

Minimize $\sum_{t=1}^{T} |y_t - x_t\beta|$, subject to $y_t \leqslant x_t\beta$, $t = 1, 2, \ldots, T$.

facilities, and an interest charge (8·12 per cent) on the closing depreciated capital value of the existing water supply. The value of the 'fencing variable', x_2, is calculated by summing the value spent on improvements to fencing and yards, depreciation on existing fences and yards, and an interest charge on the closing depreciated capital values of the yards and fences. The value of the 'labour variable', x_3, is obtained from the total cost of labour (including imputed family labour costs) minus the cost of shearing and crutching which is considered a 'marketing cost' for wool that is subtracted from gross wool sales to obtain the value of the dependent variable. The value of the 'machinery variable', x_4, is obtained by adding the costs of electricity, fuels, oils, grease, repairs to machinery, insurance and depreciation on farm plant and machinery. The value of the 'land variable', x_5, is obtained by summing the costs of rates and taxes, rents and an interest charge on the total value of land.

The production frontier of (19) has independent variables that are practically identical to those of Duloy [6, p. 87ff]. Duloy's dependent variable was net farm income and his data were obtained from the 1954-55 Sheep Industry Survey by the Bureau of Agricultural Economics. Duloy used ordinary-least squares to estimate the parameters of the production function for each of the three state regions of the Pastoral Zone of Eastern Australia. We estimate the parameters of the model (19) for the whole Pastoral Zone of Eastern Australia in addition to the three State regions of the zone. Data from 146 sample farms were used in the empirical analysis, 57 being from New South Wales, 60 from Queensland and 29 from South Australia.

Table 1 presents the maximum-likelihood estimates for the parameters of the average-, pseudo- and full-frontier production functions.[2] The logarithms of the likelihood functions for the three models are also evaluated at the respective estimates. The coefficients of determination for the ordinary least-squares regressions for the N.S.W., Queensland, S.A. and the whole zone were 0·59, 0·31, 0·86 and 0·44, respectively.

The estimates for the constant, β_0, in the pseudo-frontier model exceed the estimates obtained from the average-frontier model. In two of the three State regions (N.S.W. and Queensland) the estimates for the constant in the full-frontier model are less than the estimates for the constant in the average-frontier model. Under the assumptions of the pseudo- and full-frontier models, the ordinary least-squares estimator for β_0 has a negative bias of $(2\gamma\sigma^2/\pi)^{\frac{1}{2}}$. For the sample input values the predicted values, $x_i\hat{\beta}$, for the pseudo- and full-frontier models are never less than the observed output values or the predicted values for the average-frontier model. In some cases the predicted values for the pseudo-frontier model exceed the corresponding predicted values for the full-frontier model.

A particularly significant feature of the empirical results is the large value of the variance-ratio parameter, γ. The estimates exceed 0·95 for all regions and are significantly greater than zero. The estimates for γ are not significantly smaller than one for the three State regions and the

[2] The parameter estimates for the full-frontier model applied to the whole Eastern Australian Pastoral Zone were not obtained because the number of linear restrictions exceeded the maximum for the quadratic programming algorithm that was available.

TABLE 1

Maximum-Likelihood Estimates for the Parameters of the Production Frontiers[a]

Region	Model	Coefficients of Independent Variables						Variance Coeff.		Log-likelihood
		Constant	Water	Fencing	Labour	Machinery	Land	σ^2	γ	
New South Wales ($T = 57$)	Average	4·9099 (1·058)	0·1824 (0·114)	0·2617 (0·120)	0·1379 (0·092)	-0·0318 (0·131)	0·2041 (0·092)	0·1979 (0·037)	0·0	-34·71
	Pseudo	5·2838 (1·223)	0·2153 (0·070)	0·1167 (0·079)	0·0839 (0·075)	0·0505 (0·164)	0·2958 (0·055)	0·5047 (0·142)	0·9903 (0·061)	-30·04
	Full	4·8626	0·2446	0·1116	0·0742	0·0923	0·2936	0·5534	1·0	-24·51
Queensland ($T = 60$)	Average	3·4424 (1·983)	0·0199 (0·119)	0·6033 (0·240)	0·0816 (0·191)	0·3996 (0·292)	-0·2126 (0·234)	0·7016 (0·128)	0·0	-74·51
	Pseudo	3·7813 (1·869)	0·0543 (0·071)	0·2879 (0·197)	0·0986 (0·170)	0·3071 (0·262)	0·1901 (0·221)	1·7219 (0·376)	0·9585 (0·028)	-68·43
	Full	2·5448	0·0587	0·0	-0·0512	0·3931	0·6717	2·4331	1·0	-70·22
South Australia ($T = 29$)	Average	1·8520 (1·340)	0·2152 (0·137)	-0·0779 (0·222)	0·1073 (0·208)	-0·1217 (0·193)	0·9199 (0·235)	0·2157 (0·057)	0·0	-18·91
	Pseudo	3·6527 (1·622)	0·2446 (0·093)	-0·0034 (0·217)	0·0596 (0·184)	-0·1691 (0·158)	0·7836 (0·250)	0·5193 (0·171)	0·9868 (0·142)	-16·08
	Full	4·0176	0·3138	-0·0166	0·1145	-0·1724	0·6469	0·5502	1·0	-12·38
Eastern Australian Pastoral Zone ($T = 146$)	Average	3·8349 (0·923)	0·1373 (0·073)	0·3031 (0·120)	0·0881 (0·093)	0·0769 (0·130)	0·2471 (0·102)	0·5145 (0·060)	0·0	-158·65
	Pseudo	5·0721 (0·649)	0·1129 (0·044)	0·2308 (0·089)	0·0986 (0·061)	0·0304 (0·106)	0·3187 (0·076)	1·1518 (0·162)	0·9583 (0·018)	-136·87

[a] Standard errors of the estimators are given below the estimates.

values of the loglikelihood function indicate that the full-frontier model fits the data best in New South Wales and South Australia. Although the full-frontier model for the whole zone was not estimated, the estimated value of γ of $0 \cdot 9583$ is significantly less than one at the 5 per cent level.

The elasticities estimated for the pseudo-frontier model exceed zero with only two exceptions. In two of the four cases in which elasticities were estimated to be negative for the average-frontier model, positive estimates were obtained for the pseudo-frontier model. A comparison with the results obtained by Duloy [6] indicates that the elasticities have generally decreased over the last twenty years.

A parameter of some interest in the analysis of the production frontier (19) is the degree of homogeneity. This parameter is defined by the sum of the five coefficients $\beta_1, \beta_2, \ldots, \beta_5$. The estimates for the sum $\sum_{i=1}^{5} \beta_i$, are given in Table 2. The estimates obtained from the 1973-74 Grazing Industry Survey indicate that the production frontier does not exhibit increasing returns to scale. The standard errors for the estimators for the homogeneity parameter are such that hypotheses of constant returns to scale for the individual State regions would not be rejected if the sizes of the tests were sufficiently small. The degree of homogeneity of the production function in 1973-74 appears, however, to be significantly different to that in 1954-55.

It should be noted that in the above empirical work we are not necessarily prepared to defend our model against other competitors that suggest additional independent variables be used. Further, we are not recommending the pooling of data from diverse production units. The empirical analyses are largely numerical illustrations of the statistical model under investigation using a production function similar to the one of an earlier study.

Conclusions

In this paper we have considered a statistical model for output observations that is consistent with the traditional definition of a production function. The empirical results obtained in the estimation of sheep production functions for the Pastoral Zone of Eastern Australia indicate that the variance of asymmetric error in the model is a highly significant component. This is in contrast to the empirical results obtained by Aigner, Knox Lovell and Schmidt [3] with manufacturing and agricultural data. The latter authors found that the asymmetric error component of the error model (2) was effectively zero.

The error model (2) considered in this paper is worthy of consideration in production function analyses since it includes the familiar average model and a full-frontier model as limiting cases. Unless there are strong *a priori* reasons for assuming one particular model it is suggested that empirical data be used to determine the model that fits best.

It may often occur that a multiplicative production function is defined and there is interest in determining economic optima. In such investigations moments of the dependent variable are required. These results are

TABLE 2
Estimates for the Degree of Homogeneity of the Production Frontiers[a]

| Region | Frontier Models | | | |
| | 1973–74 Data | | | 1954–55 Data[b] |
	Average	Pseudo	Full	Average
New South Wales	0·7543 (0·127)	0·7622 (0·154)	0·8163	1·286
Queensland	0·8918 (0·228)	0·9383 (0·218)	1·0723	1·340
South Australia	1·0428 (0·147)	0·9153 (0·163)	0·8862	1·274
Eastern Australian Pastoral Zone	0·9406 (0·106)	0·7914 (0·075)	—	—

[a] Standard errors of the estimators are given below the estimates.
[b] Estimates are obtained from Duloy [6], p. 94.

presented in the Appendix for the case in which the logarithm of the production function is the model defined by equations (1) and (2).

Appendix

Consider the production function that is defined by

$$Z_t = e^{\beta_0} w_{1t}^{\beta_1} w_{2t}^{\beta_2} e^{Et}, \quad t = 1, 2, \ldots, T, \tag{A.1}$$

such that

$$E_t = U_t + V_t, \quad t = 1, 2, \ldots, T, \tag{A.2}$$

where Z_t, $t = 1, 2, \ldots, T$, are observable output values;
w_{1t} and w_{2t} are two observable input values;
β_0, β_1 and β_2 are unknown parameters; and
U_t and V_t have the same distributions as the errors in (2).

Theorem A.1: If Z_t has the distribution specified by (A.1) and (A.2) then

$$E(Z_t) = (e^{\beta_0} w_{1t}^{\beta_1} w_{2t}^{\beta_2} e^{\frac{1}{2}\sigma^2}) 2\Phi(-\sigma_U) \equiv \mu_t \tag{A.3}$$

and

$$\text{Var}(Z_t) = \mu_t^2 \{2\Phi(-2\sigma_U)e^{\sigma^2} - 4[\Phi(-\sigma_U)]^2\} \tag{A.4}$$

where $\sigma^2 = \sigma_U^2 + \sigma_V^2$.

Proof: The results (A.3) and (A.4) follow readily by well-known results dealing with expectations of functions of independent random variables and the following lemma.

Lemma A.1. If U is a negative truncated normal random variable with density function

$$f_U(u) = 2(2\pi\sigma_U^2)^{\frac{1}{2}} \exp(-\tfrac{1}{2}u^2/\sigma_U^2), \quad \text{if } u < 0$$
$$= 0, \quad \text{if } u \geq 0,$$

then the moment generating function for U is defined by

$$m_U(t) \equiv E(e^{Ut}) = 2\Phi(-\sigma_U t) \exp(\tfrac{1}{2}\sigma_U^2 t^2), \quad -\infty < t < \infty.$$

We note that if the error, e^{Et}, in (A.1) has lognormal distribution [i.e., $U_t = 0$ in (A.2) or $\sigma_U{}^2 = 0$], then the expectation and variance of Z_t are

$$E(Z_t) = (e^{\beta_0} w_{1t}{}^{\beta_1} w_{2t}{}^{\beta_2} e^{\frac{1}{2}\sigma_V^2})$$

and

$$\mathrm{Var}(Z_t) = (e^{\beta_0} w_{1t}{}^{\beta_1} w_{2t}{}^{\beta_2} e^{\frac{1}{2}\sigma_V})^2 (e^{\frac{1}{2}\sigma_V^2} - 1).$$

These results are seen to be special cases of (A.3) and (A.4) in which $\sigma_U{}^2 = 0$.

References

[1] Aigner, D. J., T. Amemiya and D. J. Poirier, 'On the Estimation of Production Frontiers'. *International Economic Review* 17: 377-396, 1976.

[2] ————— and S. F. Chu, 'On Estimating the Industry Production Function'. *American Economic Review* 58: 826-839, 1968.

[3] —————, C. A. Knox Lovell and P. Schmidt, 'Formulation and Estimation of Stochastic Frontier Production Function Models'. *Journal of Econometrics* 6(1): 21-37, 1977.

[4] Amemiya, T., 'Regression Analysis When the Dependent Variable is Truncated Normal'. *Econometrica* 41: 997-1016, 1973.

[5] Corra, G. S., *Statistical Models for Production Frontiers: With Application to the Pastoral Zone of Eastern Australia.* Unpublished B.Ag.Ec. thesis, University of New England, Armidale 1976.

[6] Duloy, J. H., *The Allocation of Resources in the Australian Sheep Industry.* Unpublished Ph.D. thesis, University of Sydney 1963.

[7] Ferguson, C. E., *Microeconomic Theory.* (Irwin: Homewood 1966.)

[8] Schmidt, P., 'On the Statistical Estimation of Parametric Frontier Production Functions'. *Review of Economics and Statistics* 58: 238-239, 1976.

[9] Theil, H., *Principles of Econometrics.* (Wiley: New York 1971.)

[10] Timmer, C. P., 'Using a Probabilistic Frontier Production Function to Measure Technical Efficiency'. *Journal of Political Economy* 79: 776-794, 1971.