University of New England

A STOCHASTIC FRONTIER PRODUCTION FUNCTION INCORPORATING A MODEL FOR TECHNICAL INEFFICIENCY EFFECTS

G.E. Battese and T.J. Coelli

No. 69 - October 1993

WORKING PAPERS IN

ECONOMETRICS AND APPLIED STATISTICS

Department of Econometrics

A STOCHASTIC FRONTIER PRODUCTION FUNCTION INCORPORATING A MODEL FOR TECHNICAL INEFFICIENCY EFFECTS*

G.E. Battese & T.J. Coelli

Department of Econometrics
University of New England
Armidale, NSW, 2351

Telephone: (067)73 2795

Draft: 13 September 1993

ABSTRACT

A stochastic frontier production function is defined for panel data on firms, in which the non-negative technical inefficiency effects are assumed to be a function of firm-specific variables and vary over time. The inefficiency effects are assumed to be independently distributed as truncations of normal distributions with constant variance, but means which are a linear function of observable firm-specific variables.

The maximum-likelihood method is applied for the estimation of the parameters of the model and the prediction of the technical efficiencies of the firms over time. The generalized likelihood-ratio test is considered for testing the null hypotheses, that the inefficiency effects are not stochastic or that they do not depend on the firm-specific variables.

An empirical application of the inefficiency stochastic frontier model is obtained using up to ten years of data on paddy farmers from an Indian village.

---

* This paper is a revision of that presented at the Australasian Meeting of the Econometric Society at the University of Sydney, 7-9 July 1993. We would like to thank Manolito Bernabe for assistance with data compilation. We gratefully acknowledge the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) for providing the panel data from its Village Level Studies.

1. Introduction

Since the stochastic frontier production function was independently proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977), there has been considerable research to extend and apply the basic frontier model. Reviews of much of this research are provied by Førsund, Lovell and Schmidt (1980), Schmidt (1986), Bauer (1990) and Battese (1992). Extensive bibliographies of empirical applications of frontier functions and efficiency analysis are given by Ley (1990) and Beck (1991).

The stochastic frontier production function postulates the existence of technical inefficiencies of production of firms involved in producing a particular output. For a given combination of input levels, it is assumed that the realized production of a firm is bounded above by the sum of a parametric function of known inputs, involving unknown parameters, and a random error, associated with measurement error of the level of production or other factors, such as the effects of weather, strikes, damaged product, etc. The greater the amount by which the realized production falls short of this stochastic frontier production, the greater the level of technical inefficiency.

Most theoretical stochastic frontier functions have not explicitly formulated a model for the inefficiency effects. Empirical papers, in which the issue of the explanation of the inefficiency effects has been raised, include Pitt and Lee (1981), Kalirajan (1981, 1982, 1989), Kalirajan and Flinn (1983) and Kalirajan and Shand (1989). These papers adopt a two-stage approach, in which the first stage involves the specification and estimation of the stochastic frontier production function and the prediction of either the inefficiency effects or the technical efficiencies of the firms involved. The second stage of the analysis involves the specification of a regression model for either the predicted inefficiency effects or the levels of technical efficiency of the firms in terms of various explanatory variables and an additive random error. The parameters of this second-stage inefficiency model have been generally estimated by using ordinary least-squares regression. Kalirajan (1981) specifies that the random errors in the second-stage inefficiency model have half-normal distribution. In all these empirical studies, the methods of estimation of the parameters of the second-stage inefficiency model are based on assumptions which are clearly false, because the effects of estimation of the stochastic frontier model were not accounted for.[1]

Pitt and Lee (1981) investigate the sources of technical inefficiency by specifying that firm intercepts in the stochastic frontier are a function of firm characteristics. The authors regress the estimated firm intercepts on the specified firm characteristics or incorporate the firm characteristics into the production frontier and jointly estimate the parameters involved.

More recently, models for the inefficiency effects in stochastic frontier production functions have been proposed by Kumbhakar, Ghosh and McGuckin (1991), Reifschneider and Stevenson (1991) and Huang and Liu (1992). Kumbhakar, Ghosh and McGuckin (1991) assume that the technical inefficiency effects are non-negative truncations of a normal distribution with mean, which is a linear function of exogeneous factors whose coefficients are unknown, and an unknown variance. In addition, Kumbhakar, Ghosh and McGuckin (1991) consider allocative inefficiencies associated with the side conditions for profit maximization not being exactly satisfied. In the

---

[1] For example, many studies assume the firm effects (usually denoted by U_i) are independently distributed in the first-stage estimation. They then regress the predicted U_i's upon firm specific factors in a second stage. The specification of this second-stage model clearly conflicts with the assumption that the U_i are independent.

application of their model to US dairy farms, Kumbhakar, Ghosh and McGuckin (1991) find that the technical inefficiency effects are significantly related to the level of education of the farmers and the size of their farming operations. Technical and allocative inefficiencies are investigated in the context of a frontier production function of Zellner-Revankar (1969) type which proves to be significantly different from the Cobb-Douglas model.

Reifschneider and Stevenson (1991) propose a model for the inefficiency effects of the stochastic frontier production function involving the sum of a non-negative function of relevant explanatory variables and a non-negative random variable, which is assumed to have half-normal, exponential or gamma distribution. This model is applied in the analysis of data on electricity generation in the USA during three different time periods. The hypothesis, that the inclusion of the inefficiency function does not change the estimates of the frontier function parameters, is rejected in their study.

Huang and Liu (1992) consider a stochastic frontier production function in which the non-negative technical inefficiency effects are a linear function of variables involving firm characteristics. The additive random error of the inefficiency model is assumed to be the truncation of a normal distribution with mode zero, whose point of truncation is dependent on the firm characteristics, such that the inefficiency effects are non-negative. Hence the random errors are not required to be non-negative, as in the model of Reifschneider and Stevenson (1991). Huang and Liu (1992) apply their inefficiency frontier model in the analysis of cross-sectional data from the electronics industry in Taiwan and assume that the explanatory variables in the inefficiency model are a function of firm-specific variables and the explanatory variables of the stochastic frontier. This makes their model a non-neutral shift of the traditional average response function, in that the marginal products of inputs and marginal rates of technical substitution depend on the firm-specific variables in the inefficiency model.

The present paper extends the Huang and Liu (1992) model to panel data. This allows us to include both firm-specific effects and time effects in the model of inefficiency. The model is applied to farm-level data from an Indian village.

2. Inefficiency Frontier Model for Panel Data

Consider the stochastic frontier production function for panel data, which is defined by equation (1),

$$Y_{it} = \exp(x_{it}\beta + V_{it} - U_{it}) \tag{1}$$

where Y_{it} denotes the production for the t-th observation (t = 1,2,...,T) for the i-th firm (i = 1,2,...,N);

x_{it} is a (1×k) vector of values of known functions of inputs of production associated with the i-th firm at the t-th period of observation;

β is a (k×1) vector of unknown parameters to be estimated;

the V_{it}'s are assumed to be iid $N(0, \sigma_V^2)$ random errors, independently distributed of the U_{it}'s which are non-negative random variables, associated with technical inefficiency of production;

the U_{it}'s are assumed to be independently distributed, such that U_{it} is obtained by truncation (at zero) of the normal distribution with mean, $z_{it}\delta$, and variance, σ^2;

z_{it} is a (1×m) vector of firm-specific variables which may vary over time; and

δ is an (m×1) vector of unknown coefficients of the firm-specific inefficiency variables.

Although it is assumed that there are T time periods for which observations are available for at least one of the N firms involved, it is not necessary that all the firms are observed for all T periods.

Equation (1) specifies the stochastic frontier production function (e.g., of Cobb-Douglas or transcendental-logarithmic form) in terms of the original production values. However, the technical inefficiency effects, the U_{it}'s, are assumed to be a function of a set of explanatory variables, the z_{it}'s, and an unknown vector of coefficients, δ. The explanatory variables in the inefficiency model would be expected to include any variables which explain the extent to which the production observations fall short of the corresponding stochastic frontier production values, $\exp(x_{it}\beta + V_{it})$. The z_{it}-vectors may have the first element equal to one, include some input variables involved in the production function and/or interactions between firm-specific variables and input variables. If the first z-variable has value one and the coefficients of all other z-variables are zero, then this case would represent the model specified by Stevenson (1980) and Battese and Coelli (1988, 1992). If all elements of the δ-vector were equal to zero, then the inefficiency effects are not related to the z-variables and so the half-normal distribution originally specified by Aigner, Lovell and Schmidt (1977) would be obtained. If interactions between firm-specific variables and input variables are included, then the non-neutral model proposed by Huang and Liu (1993) is obtained.

The inefficiency effects, U_{it}, in the stochastic frontier model (1) could be specified in equation (2),

$$U_{it} = z_{it}\delta + W_{it} \qquad (2)$$

where the random variable, W_{it}, is defined by the truncation of the normal distribution with zero mean and variance, σ^2, such that the point of truncation is $-z_{it}\delta$, i.e., $W_{it} \geq -z_{it}\delta$. These assumptions are consistent with the U_{it}'s being non-negative truncations of the $N(z_{it}\delta, \sigma^2)$-distribution.

The assumption that the U_{it}'s are independently distributed for all $t = 1,2,\ldots,T$, and $i = 1,2,\ldots,N$, is obviously a simplifying, but restrictive, condition. Alternative models are required to account for possible correlated structures of the inefficiency effects over time.

It should be noted that the inefficiency frontier model (1)-(2) is not a generalization of the Battese and Coelli (1992) model for time-varying inefficiencies, even if the inefficiency effects are time invariant. The Battese and Coelli (1992) model specifies that the inefficiency effects are the product of an exponential function of time and non-negative firm-specific random variables, i.e., $U_{it} = \left\{\exp[-\eta(t-T)]\right\}U_i$, where η is an unknown parameter and the U_i's are non-negative truncations of the $N(\mu, \sigma^2)$ distribution. This model does not define the inefficiency effects in terms of firm-specific explanatory variables. Further, the Battese and Coelli (1992) model specifies well-defined correlated structures for the inefficiency effects over time for particular firms.

When the model in equation (1) is assumed, the technical efficiency of production for the i-th firm at the t-th observation is defined by equation (3).

$$TE_{it} = \exp(-U_{it}) = \exp(-z_{it}\delta - W_{it}) \tag{3}$$

Since $z_{it}\delta + W_{it} > z_{i't}\delta + W_{i't}$ for $i \neq i'$ does not necessarily imply that $z_{it'}\delta + W_{it'} > z_{i't'}\delta + W_{i't'}$ for $t' \neq t$, then it follows that the same ordering of firms in terms of technical efficiency of production does not apply for all time periods, as for the Battese and Coelli (1992) model.

The inefficiency frontier production function (1)-(2) differs from that of Reifschneider and Stevenson (1991) in that the W-random variables are not identically distributed, as in the latter paper. Reifschneider and Stevenson (1991) assume that the W-random variables in the inefficiency model are non-negative random variables which have half-normal, exponential or gamma distribution. In our model, the W-random variables could be negative if $z_{it}\delta > 0$, i.e., $W_{it} \geq -z_{it}\delta$, but are independent truncations of the normal distribution with zero mean and variance, σ^2.

The inefficiency model (1)-(2) is closely related to the Huang and Liu (1992) model, in that it is a straightforward extension of the latter model to account for panel data. This extension for time-series data has the same distributional assumptions as if the cross-sectional dimension of the data was increased. However, for our panel-data model there would be particular interest in the behaviour of the technical efficiencies of production of the panel of firms over time.

The parameters of the model defined by (1) and (2) may be estimated by the method of maximum likelihood. The derivation of the likelihood function and its partial derivatives with respect to the parameters of the model are presented in the Appendix.

3. Empirical Application

Data on paddy farmers from the Indian village of Aurepalle are considered for an empirical application of the inefficiency stochastic frontier production function discussed in the previous section. These data were collected by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Battese and Coelli (1992) use data on 15 farmers over the ten-year period from 1975-76 to 1984-85. Because 21 observations were not available for some farmers in some of the years involved, only 129 observations were used in that paper.

We have endeavoured to obtain data on farmer characteristics that may influence technical efficiency. Information on the age and years of schooling of 14 of the 15 farmers was obtained. Hence the data in the present study comprise 14 farmers and a total of 125 observations. Information on other variables, such as the frequency of contacts with agricultural extension officers, access to credit and the use of high-yielding varieties, fertilizers, etc., were not readily available. While not providing a thorough analysis, we believe that the use of age, years of formal schooling and year of observation satisfactorily illustrate the methodology involved.

The stochastic frontier production function to be estimated is defined by equation (4),

$$\log(Y_{it}) = \beta_0 + \beta_1 \log(Land_{it}) + \beta_2 (PILand_{it}) + \beta_3 \log(Labour_{it}) + \beta_4 \log(Bullocks_{it}) + \beta_5 \log[Max(Costs_{it}, 1-D_{it})] + \beta_6 (Year_{it}) + V_{it} - U_{it} \quad (4)$$

where the technical inefficiency effects are assumed to be defined by equation (5),

$$U_{it} = \delta_0 + \delta_1(Age_{it}) + \delta_2(Schooling_{it}) + \delta_3(Year_{it}) + W_{it} \quad (5)$$

where

Y is the total value of output (in Rupees);

Land is the total area of irrigated and unirrigated land operated (in hectares);

PILand is the proportion of the operated land that is irrigated;

Labour is the total amount of family and hired labour engaged in farming (in male equivalent hours);

Bullocks represents the amount of bullock labour used (in hours of bullock pairs);

Costs refers to the total value of all other inputs used (fertilizer, manure, pesticides, machinery, etc.);

D is a dummy variable which has value one if *Costs* are positive and zero, otherwise;

Age is the age of the primary decision maker in the farming operation;

Schooling refers to the number of years of formal schooling of the primary decision maker;

Year indicates the year of the observation involved; and

V_{it} and W_{it} are as defined in the previous section.

The model in (4) can be viewed as a linearized version of the logarithm of the Cobb-Douglas production function in which the land variable is a weighted average of the number of irrigated and unirrigated hectares of land used in the production of paddy and other crops. The variable, *PILand*, accounts for the differences in the productivities of irrigated and unirrigated land. More details on the functional form of the land and labour variables involved in the production frontier (4) are given by Battese, Coelli and Colby (1989) and Battese and Coelli (1992).

The variables in the production function (4) are those which are in the preferred model in Battese and Coelli (1992). However, the stochastic properties of that model are identical to the ordinary-least squares model, given the assumptions of the stochastic frontier model with time-varying inefficiencies proposed by Battese and Coelli (1992). In this paper, however, inefficiency effects are assumed to be present in the stochastic frontier and be linearly related to age and education of the paddy farmers and the year of observation involved, such that an intercept parameter is included.

The inefficiency frontier model, defined by equations (4)-(5), account for both technical change and time-varying technical inefficiency effects. The *Year* variable in the stochastic frontier production function (4) accounts for Hicksian neutral technological change. However, the *Year* variable in the inefficiency model (5) specifies that the inefficiency effects may change linearly with respect to time. The distributional assumptions on the inefficiency effects permit the effects of technical change and time-varying technical inefficiencies to be identified, in addition to the intercept parameters in the stochastic frontier and the inefficiency model, given the specifications of the time effects involved.

Maximum-likelihood estimates of the parameters of the model, defined in (4)-(5), were obtained using FRONTIER 3.0, a modification of the computer program, FRONTIER 2.0 (see Coelli, 1992), which was written to estimate the time-varying inefficiency frontier model of Battese and Coelli (1992). These results are given in the second last column of Table 1, indicated by Model 1. The last column of Table 1 gives the maximum-likelihood estimates for the parameters of the preferred frontier model, to be discussed below, in which some parameters in the general model are specified to be zero.

The signs of the β-estimates are all as expected, with the exception of the negative estimate of the bullock-labour variable. Possible reasons for the negative parameter associated with bullock labour are discussed by Saini (1979), Battese, Coelli and Colby (1989) and Battese and Coelli (1992). The positive coefficient of the proportion of land which is irrigated confirms the expected positive relationship between the proportion of irrigated land and total production.

The coefficients of the explanatory variables in the inefficiency model (5) are of particular interest to this study. The estimate for the coefficient associated with *Age* is positive, which indicates that the older farmers are more technically inefficient as paddy farmers than the younger ones. The estimate for the coefficient associated with *Schooling* is negative. This implies that the paddy farmers with greater years of schooling tend to be more technically efficient. However, the relationship is very weak, because the coefficient is highly insignificant (by an asymptotic t-test[2]).

---

2 Unless otherwise stated, all tests of hypotheses this paper are conducted at the 5% level of significance.

Table 1: Maximum-likelihood Estimates for Parameters of Selected Inefficiency Stochastic Frontier Production Functions for Paddy Farmers in Aurepalle*

Variable	Parameter	Model 1	Model 2
Constant	β_0	2.86 (0.60)	3.01 (0.57)
log(Land)	β_1	0.37 (0.12)	0.37 (0.13)
PILand	β_2	0.38 (0.21)	0.42 (0.23)
log(Labour)	β_3	0.85 (0.13)	0.79 (0.12)
log(Bullocks)	β_4	-0.33 (0.11)	-0.28 (0.10)
log(Costs)	β_5	0.071 (0.031)	0.084 (0.032)
Year	β_6	0.014 (0.013)	0
Constant	δ_0	-1.5 (2.8)	0
Age	δ_1	0.035 (0.034)	0.0154 (0.0046)
Schooling	δ_2	-0.006 (0.077)	0
Year	δ_3	-0.57 (0.60)	-0.34 (0.20)
	σ_S^2	0.74 (0.75)	0.40 (0.20)
	γ	0.952 (0.047)	0.922 (0.048)
Log (likelihood)		-22.595	-23.057

* Estimated standard errors are given in parentheses to two significant digits. The estimated coefficients are given to the corresponding numbers of digits behind the decimal places.

The negative coefficient of the *Year* variable suggests that technical inefficiency of production of the paddy farmers declined throughout the ten-year period.

The remaining two parameters, $\sigma_S^2 \equiv \sigma_V^2 + \sigma^2$ and $\gamma \equiv \sigma^2/\sigma_S^2$, are associated with the variances of the random variables, V_{it} and U_{it}. Generalized likelihood-ratio tests[3] of the hypotheses that the technical inefficiency effects are absent or that they have simpler distributions are presented in Table 2.

Table 2: Tests of Hypotheses for Parameters of the Inefficiency Stochastic Frontier Production Function for Paddy Farmers in Aurepalle

Null Hypotheses	Log(Likelihood)	χ^2-statistic	Decision
H_0: $\gamma = \delta_0 = \ldots = \delta_3 = 0$	-37.588	29.99	Reject H_0
H_0: $\gamma = \delta_0 = \delta_3 = 0$	-36.082	26.97	Reject H_0
H_0: $\delta_1 = \delta_2 = \delta_3 = 0$	-27.941	10.69	Reject H_0
H_0: $\delta_0 = 0$	-22.892	0.59	Accept H_0
H_0: $\beta_6 = \delta_0 = \delta_2 = 0$	-23.057	0.92	Accept H_0
Restriction: $\beta_6 = \delta_0 = \delta_2 = 0$			
H_0: $\gamma = 0$	-36.251	26.39	Reject H_0
H_0: $\delta_1 = \delta_3 = 0$	-39.021	31.93	Reject H_0

3 The likelihood-ratio test statistic is calculated as

$$\lambda = -2[\log(\text{likelihood }(H_0)) - \log(\text{likelihood }(H_1))]$$

and has approximately chi-square distribution with parameter equal to the number of parameters assumed to be equal to zero in the null hypothesis, H_0.

The null hypothesis that the inefficiency effects are absent from the model (i.e., H_0: $\gamma = \delta_0 = \delta_1 = \delta_2 = \delta_3 = 0$) is rejected. The second null hypothesis considered in Table 2, H_0: $\gamma = \delta_0 = \delta_3 = 0$, specifies that the inefficiency effects are not stochastic. If the parameter, γ, is zero, then the variance of the inefficiency effects is zero and so the model reduces to a traditional mean response function in which the variables, age and schooling of the farmers, are included in the production function. The parameters, δ_0 and δ_3, must be zero if γ is zero, given that the production function involves an intercept parameter and year of observation. If there are no random inefficiency effects in the model, then the parameters, δ_0 and δ_3, are not identified. However, the null hypothesis that the inefficiency effects are not random is rejected.

The null hypothesis that the inefficiency effects are not a linear function of the year of observation and the age and schooling of the farm operator, H_0: $\delta_1 = \delta_2 = \delta_3 = 0$, is also rejected. This indicates that the joint effects of these three explanatory variables on the levels of technical inefficiencies is significant, although the individual effects of one or more of the variables may not be statistically significant. However, the hypothesis that the inefficiency effects do not have an intercept parameter is not rejected.

Because the estimate for the intercept parameter in the inefficiency model is small relative to its estimated standard error, the model was re-estimated without this parameter. As expected, the estimates for the parameters in this model were little different from those obtained for the more general model, but the estimated coefficients of year of observation in the frontier and schooling in the inefficiency model (β_6 and δ_2, respectively) were less than their estimated standard errors. In fact, the chi-square statistic for testing the null hypothesis, H_0: $\beta_6 = \delta_0 = \delta_2 = 0$, is not significant and so we consider that the preferred frontier model has the three parameters, β_6, δ_0 and δ_2, equal to zero.

The maximum-likelihood estimates for the parameters of the preferred frontier model are presented in the last column of Table 1. All the parameter estimates for this model are considerably larger than their estimated standard errors. The chi-square statistics for testing the null hypotheses of the absence of stochastic inefficiency effects (H_0: $\gamma = 0$) and of the absence of age and year effects in the inefficiency model (H_0: $\delta_1 = \delta_3 = 0$) in the preferred frontier model are not significant (see Table 2).

The parameter estimates for the preferred stochastic frontier production function indicate that the elasticity of land is estimated to be 0.37. The estimated elasticity for labour, 0.79, is quite large. The elasticity for bullock labour is significantly less than zero. The estimated elasticity for other input costs is relatively small, 0.084, but is significantly different from zero. These estimates imply that the returns-to-scale parameter is estimated to be 0.965, with estimated standard error of 0.048. Thus the technology of the paddy farmers is such that the hypothesis of constant returns to scale would be accepted.

The technical inefficiency effects in the preferred model are significant, such that older farmers tend to have greater values of technical inefficiency. However, the levels of technical inefficiency for the paddy farmers tend to decrease over time.

The technical efficiencies of the paddy farmers in the different years involved can be obtained using the predictor, presented in equation (A.10) of the Appendix. The parameters involved are estimated by their maximum-likelihood estimates. The predicted technical efficiencies obtained for the 14 paddy farmers involved are presented in Table 3.

The predicted technical efficiencies show considerable variability among the paddy farmers. The technical efficiencies of individual paddy farmers also vary up and down over time. Some farmers had the highest level of technical efficiency in one or more years, but had the lowest technical efficiency in at least one year, as well. For example, Farmer 1 had the highest technical efficiencies in the years 1975-76 and 1977-78, but also had

Table 3: Technical Efficiencies of Paddy Farmers in Aurepalle

Farmer	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85
1	.887	.615	.928	.606	.856	.730	.733	.944	.839	.814
2	.724	.628	.898	.622	.853	.712	.727	.944	.835	.873
3	.518	.215	.835	.847	.908	.653	-	-	-	-
4	.540	.287	.751	.902	.777	.565	.744	.876	.872	.918
5	.460	.606	.886	.768	.837	.778	.904	.838	.918	.852
6	.730	.510	.922	.866	.794	.767	.761	.913	.884	.885
7	.505	.310	.914	.824	.759	.715	.763	.899	.797	-
8	.758	.465	.749	-	.873	.690	.906	.936	.899	.908
9	.623	.229	.792	.793	.820	.742	.763	.901	.904	.928
10	.664	.737	.875	.812	.913	.723	.928	.940	.898	.943
11	.718	.399	.819	.819	.868	.755	.885	.904	.861	.941
12	.569	.486	.886	.799	.764	.753	.915	.946	.914	.937
13	.420	.402	.888	.800	.825	.293	.570	-	-	-
14	-	.410	.884	.861	.896	-	-	-	-	-

the lowest technical efficiencies among the paddy farmers in 1978-79 and 1984-85. These values indicate that there is considerable variation in the levels of technical efficiencies over time for given paddy farmers, although there is a general decline in the technical inefficiency of the paddy farmers over time.

4. Conclusions

The results obtained in the empirical application of the proposed inefficiency stochastic frontier production function exhibit some interesting differences from those obtained in the application of the time-varying inefficiency model presented by Battese and Coelli (1992). With the latter model, it was concluded that there was no evidence of technical inefficiencies of production with essentially the same sample of paddy farmers as in this study. However, the Battese and Coelli (1992) model assumed that the technical inefficiency effects were the product of an exponential function of time and non-negative firm-specific random variables. The present model specifies that the inefficiency effects are a linear function of some firm-specific variables and time, together with an additive stochastic error which is assumed to be independent over time and among firms. The two models involved are clearly separate and so it is difficult to conclude which is the "best" model for the data involved. However, the logarithm of the likelihood function for the data is greater under the assumptions of the above inefficiency stochastic frontier model than for the Battese and Coelli (1992) model.

Further theoretical and applied work is obviously required to obtain better and more general models for stochastic frontiers and the inefficiency effects involved.

References

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", **Journal of Econometrics** 6, 21-37.

Battese, G.E. (1992), "Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics", **Agricultural Economics** 7, 185-208.

Battese, G.E. and T.J. Coelli (1988), "Prediction of Firm-Level Technical Efficiencies With a Generalized Frontier Production Function and Panel Data", **Journal of Econometrics** 38, 387-399.

Battese, G.E. and T.J. Coelli (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", **Journal of Productivity Analysis** 3, 153-169.

Battese, G.E., T.J. Coelli and T.C. Colby (1989), "Estimation of Frontier Production Functions and the Efficiencies of Indian Farms Using Panel Data From ICRISAT's Village Level Studies", **Journal of Quantitative Economics** 5, 327-348.

Bauer, P.W. (1990), "Recent Developments in the Econometric Estimation of Frontiers", **Journal of Econometrics** 46, 39-56.

Beck, M. (1991), "Empirical Applications of Frontier Functions: A Bibliography", mimeo, Joachim-Ringelnatz-Str.20, W-6200 Wiesbaden, Germany, pp.9.

Coelli, T.J. (1992), "A Computer Program for Frontier Production Function Estimation: FRONTIER, Version 2.0", **Economics Letters** 39, 29-32.

Førsund, F.R., C.A.K. Lovell and P. Schmidt (1980), "A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement", **Journal of Econometrics** 13, 5-25.

Huang, C. and J-T. Liu (1992), "Stochastic Production Frontier in the Taiwan Electronics Industry", unpublished paper, Department of Economics, Vanderbilt University, Nashville, pp.13.

Kalirajan, K. (1981), "An Econometric Analysis of Yield Variability in Paddy Production", **Canadian Journal of Agricultural Economics** 29, 283-294.

Kalirajan, K. (1982), "On Measuring Yield Potential of the High Yielding Varieties Technology at Farm Level", **Journal of Agricultural Economics** 33, 227-236.

Kalirajan, K.P. (1989), "On Measuring the Contribution of Human Capital to Agricultural Production", **Indian Economic Review** 24, 247-261.

Kalirajan, K.P. and J.C. Flinn (1983), "The Measurement of Farm-specific Technical Efficiency", **Pakistan Journal of Applied Economics** 2, 167-180.

Kalirajan, K. and R.T. Shand (1986), "Estimating Location-specific and Firm-specific Technical Efficiency: An Analysis of Malaysian Agriculture", **Journal of Economic Development** 11, 147-160.

Kumbhakar, S.C., S. Ghosh and J.T. McGuckin (1991), "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms", **Journal of Business and Economic Statistics** 9, 279-286.

Ley, E. (1990), "A Bibliography on Production and Efficiency", mimeo, Department of Economics, University of Michigan, Ann Arbor, MI, pp.32.

Meeusen, W. and J. van den Broeck (1977), "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error", **International Economic Review** 18, 435-444.

Reifschneider, D. and R. Stevenson (1991), "Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency", **International Economic Review** 32, 715-723.

Saini, G.R. (1979), **Farm Size, Resource-Use Efficiency and Income Distribution**, New Delhi: Allied Publishers.

Stevenson, R.F. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation", **Journal of Econometrics** 13, 57-66.

Zellner, A. and N.S. Revankar (1969), "Generalized Production Functions", **Review of Economic Studies** 36, 241-250.

Appendix

For simplicity of presentation of results in this Appendix, we assume that the inefficiency stochastic frontier model is expressed by

$$Y_{it} = x_{it}\beta + E_{it} \tag{A.1}$$

and

$$E_{it} = V_{it} - U_{it} , \tag{A.2}$$

where $i = 1,2,\ldots,N$, and $t = 1,2,\ldots,T$.

Thus, given the frontier production function of equation (2.1), Y_{it} in equation (A.1) is, in fact, the logarithm of the production for the i-th firm in the t-th time period. Apart from introducing the notation, E_{it}, for the difference between V_{it} and U_{it}, all other variables and parameters in (A.1)-(A.2) are the same as in (1)-(2).

The density functions for V_{it} and U_{it} are

$$f_V(v) = \frac{\exp(-\frac{1}{2}\ v^2/\sigma_V^2)}{\sqrt{2\pi}\ \sigma_V} , \quad -\infty < v < \infty \tag{A.3}$$

and

$$f_U(u) = \frac{\exp[-\frac{1}{2}(u-z\delta)^2/\sigma^2]}{\sqrt{2\pi}\ \sigma\ \Phi[z\delta/\sigma]} , \quad u \geq 0 , \tag{A.4}$$

where the subscripts, i and t, are omitted for convenience in the presentation; and $\Phi(\cdot)$ represents the distribution function for the standard normal random variable.

The joint density function for $E = V-U$ and U is

$$f_{E,U}(e,u) = \frac{\exp-\frac{1}{2}\{[(e+u)^2/\sigma_V^2] + [(u-z\delta)^2/\sigma^2]\}}{2\pi\ \sigma\ \sigma_V\ \Phi[z\delta/\sigma]} , \quad u \geq 0$$

$$= \frac{\exp-\frac{1}{2}\{[(u-\mu_*)^2/\sigma_*^2] + (e^2/\sigma_V^2) + (z\delta/\sigma)^2 - (\mu_*/\sigma_*)^2\}}{2\pi\ \sigma\ \sigma_V\ \Phi[z\delta/\sigma]} \tag{A.5a}$$

or, alternatively,

$$f_{E,U}(e,u) = \frac{\exp-\frac{1}{2}\{[(u-\mu_*)^2/\sigma_*^2] + [(e+z\delta)^2/(\sigma_V^2+\sigma^2)]\}}{2\pi\ \sigma\ \sigma_V\ \Phi(z\delta/\sigma)} \tag{A.5b}$$

where

$$\mu_* = \frac{\sigma_V^2 z\delta - \sigma^2 e}{\sigma_V^2 + \sigma^2} \tag{A.6}$$

and

$$\sigma_*^2 = \sigma^2\sigma_V^2/(\sigma^2 + \sigma_V^2). \tag{A.7}$$

Thus the density function for E = V-U is

$$f_E(e) = \frac{\exp-\frac{1}{2}\{(e^2/\sigma_V^2) + (z\delta/\sigma)^2 - (\mu_*/\sigma_*)^2\}}{\sqrt{2\pi}\,\sigma_V\sigma\ \Phi[z\delta/\sigma]}\int_0^\infty \frac{\exp-\frac{1}{2}[(u-\mu_*)^2/\sigma_*^2]}{\sqrt{2\pi}}\,du$$

$$= \frac{\exp-\frac{1}{2}\{(e^2/\sigma_V^2) + (z\delta/\sigma)^2 - (\mu_*/\sigma_*)^2\}}{\sqrt{2\pi}\ (\sigma^2+\sigma_V^2)^{1/2}[\Phi(z\delta/\sigma)/\Phi(\mu_*/\sigma_*)]}\ , \tag{A.8a}$$

or, alternatively,

$$f_E(e) = \frac{\exp\{-\frac{1}{2}(e+z\delta)^2/(\sigma_V^2+\sigma^2)\}}{\sqrt{2\pi}\ (\sigma_V^2+\sigma^2)^{1/2}[\Phi(z\delta/\sigma)/\Phi(\mu_*/\sigma_*)]}\ . \tag{A.8b}$$

The conditional density function for U given E = e is thus

$$f_{U|E=e}(u) = \frac{\exp-\frac{1}{2}[(u-\mu_*)^2/\sigma_*^2]}{\sqrt{2\pi}\,\sigma_*\ \Phi(\mu_*/\sigma_*)}\ ,\quad u \geq 0\ . \tag{A.9}$$

It can be shown that the conditional expectation of e^{-U}, given E = e, is

$$E(e^{-U}|E=e) = \left\{\exp[-\mu_* + \tfrac{1}{2}\sigma_*^2]\right\}\left\{\Phi[(\mu_*/\sigma_*)-\sigma^*]/\Phi(\mu_*/\sigma_*)\right\}\ . \tag{A.10}$$

The density function for the production value, Y_{it}, in equation (A.1), is most conveniently given using the expression in equation (A.8b),

$$f_{Y_{it}}(y_{it}) = \frac{\exp-\frac{1}{2}\left\{\frac{(y_{it} - x_{it}\beta+z_{it}\delta)^2}{\sigma_V^2+\sigma^2}\right\}}{\sqrt{2\pi}\ (\sigma_V^2+\sigma^2)^{1/2}[\Phi(d_{it})/\Phi(d^*_{it})]} \tag{A.11}$$

where $d_{it} = z_{it}\delta/\sigma$, $d^*_{it} = \mu^*_{it}/\sigma_*$ and $\mu^*_{it} = [\sigma_V^2\ z_{it}\delta - \sigma^2(y_{it}-x_{it}\beta)]/(\sigma_V^2+\sigma^2)$.

Given that there are T_i observations obtained for the i-th firm, where $1 \le T_i \le T$, and $Y_i \equiv (Y_{i1}, Y_{i2}, \ldots, Y_{iT_i})'$ denotes the vector of the T_i production values in equation (A.1), then the logarithm of the likelihood function for the sample observations, $y \equiv (y_1', y_2', \ldots, y_T')'$, is

$$\begin{aligned} L^*(\theta^*;y) = & -\tfrac{1}{2}\left(\sum_{i=1}^{N} T_i\right)\left\{\ell n\ 2\pi + \ell n(\sigma^2 + \sigma_V^2)\right\} \\ & -\tfrac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[(y_{it}-x_{it}\beta+z_{it}\delta)^2/(\sigma_V^2+\sigma^2)\right] \\ & -\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\ell n\ \Phi(d_{it}) - \ell n\ \Phi(d^*_{it})\right] \end{aligned} \tag{A.12}$$

where $\theta^* = (\beta', \delta', \sigma_V^2, \sigma^2)'$.

Using the re-parameterization of the model involving the parameters, $\sigma_S^2 \equiv \sigma_V^2 + \sigma^2$ and $\gamma \equiv \sigma^2/\sigma_S^2$, the logarithm of the likelihood function is expressed by

$$\begin{aligned} L^*(\theta;y) = & -\tfrac{1}{2}\left(\sum_{i=1}^{N} T_i\right)\left\{\ell n\ 2\pi + \ell n\ \sigma_S^2\right\} \\ & -\tfrac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left\{(y_{it}-x_{it}\beta+z_{it}\delta)^2/\sigma_S^2\right\} \\ & -\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left\{\ell n\ \Phi(d_{it}) - \ell n\ \Phi(d^*_{it})\right\} \end{aligned} \tag{A.13}$$

where $d_{it} = z_{it}\delta/(\gamma\sigma_S^2)^{1/2}$ (A.14)

$$d^*_{it} = \mu^*_{it}/[\gamma(1-\gamma)\sigma^2_S]^{1/2} \quad (A.15)$$

$$\mu^*_{it} = (1-\gamma)z_{it}\delta - \gamma(y_{it}-x_{it}\beta) \quad (A.16)$$

$$\sigma_* = [\gamma(1-\gamma)\sigma^2_S]^{1/2} \quad (A.17)$$

and $\theta = (\beta', \delta', \sigma^2_S, \gamma)'$.

The partial derivatives of the logarithm of the likelihood function (A.13) with respect to the parameters, β, δ, σ^2_S and γ, are given by

$$\frac{\partial L^*}{\partial \beta} = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left\{ \frac{(y_{it}-x_{it}\beta+z_{it}\delta)}{\sigma^2_S} + \frac{\phi(d^*_{it})}{\Phi(d^*_{it})} \frac{\gamma}{\sigma_*} \right\} x'_{it}$$

where $\phi(\cdot)$ represents the density function for the standard normal random variable;

$$\frac{\partial L^*}{\partial \delta} = - \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left\{ \frac{(y_{it}-x_{it}\beta+z_{it}\delta)}{\sigma^2_S} + \left[\frac{\phi(d_{it})}{\Phi(d_{it})} \cdot \frac{1}{(\gamma\, \sigma^2_S)^{1/2}} - \frac{\phi(d^*_{it})}{\Phi(d^*_{it})} \cdot \frac{(1-\gamma)}{\sigma_*} \right] \right\} z'_{it}$$

$$\frac{\partial L^*}{\partial \sigma^2_S} = - \frac{1}{2}\left(\frac{1}{\sigma^2_S}\right) \left\{ \left(\sum_{i=1}^{N} T_i \right) - \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[\frac{\phi(d_{it})}{\Phi(d_{it})} d_{it} - \frac{\phi(d^*_{it})}{\Phi(d^*_{it})} d^*_{it} \right] \right.$$

$$\left. - \sum_{i=1}^{N} \sum_{t=1}^{T_i} \frac{(y_{it}-x_{it}\beta+z_{it}\delta)}{\sigma^2_S} \right\}$$

$$\frac{\partial L^*}{\partial \gamma} = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left\{ \frac{\phi(d_{it})}{\Phi(d_{it})} \frac{d_{it}}{2\gamma} + \frac{\phi(d^*_{it})}{\Phi(d^*_{it})} \left[\frac{y_{it}-x_{it}\beta+z_{it}\delta}{\sigma_*} + \frac{d^*_{it}(1-2\gamma)}{2\gamma(1-\gamma)\sigma^2_*} \right] \right\} .$$