The Journal of Productivity Analysis, 3, 153–169 (1992)

Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India*

G.E. BATTESE AND T.J. COELLI
Department of Econometrics, University of New England, Armidale, NSW 2351, Australia

Abstract

Frontier production functions are important for the prediction of technical efficiencies of individual firms in an industry. A stochastic frontier production function model for panel data is presented, for which the firm effects are an exponential function of time. The best predictor for the technical efficiency of an individual firm at a particular time period is presented for this time-varying model. An empirical example is presented using agricultural data for paddy farmers in a village in India.

1. Introduction

The stochastic frontier production function, which was independently proposed by Aigner, Lovell, and Schmidt [1977] and Meeusen and van den Broeck [1977], has been a significant contribution to the econometric modeling of production and the estimation of technical efficiency of firms. The stochastic frontier involved two random components, one associated with the presence of technical inefficiency and the other being a traditional random error. Prior to the introduction of this model, Aigner and Chu [1968], Timmer [1971], Afriat [1972], Richmond [1974], and Schmidt [1976] considered the estimation of deterministic frontier models whose values were defined to be greater than or equal to observed values of production for different levels of inputs in the production process.

Applications of frontier functions have involved both cross-sectional and panel data. These studies have made a number of distributional assumptions for the random variables involved and have considered various estimators for the parameters of these models. Survey papers on frontier functions have been presented by Førsund, Lovell, and Schmidt [1980], Schmidt [1986], Bauer [1990] and Battese [1992], the latter article giving particular attention to applications in agricultural economics. Beck [1991] and Ley [1990] have compiled extensive bibliographies on empirical applications of frontier functions and efficiency analysis.

The concept of the technical efficiency of firms has been pivotal for the development and application of econometric models of frontier functions. Although technical efficiency may be defined in different ways (see, e.g., Färe, Grosskopf, and Lovell [1985]), we consider

*This article is a revision of the Invited Paper presented by the senior author in the "Productivity and Efficiency Analysis" sessions at the ORSA/TIMS 30th Joint National Meeting, Philadelphia, Pennsylvania, 29–31 October 1990. We have appreciated comments from Martin Beck, Phil Dawson, Knox Lovell and three anonymous referees. We gratefully acknowledge the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) for making available to us the data obtained from the Village Level Studies in India.

the definition of the technical efficiency of a given firm (at a given time period) as the ratio of its mean production (conditional on its levels of factor inputs and firm effects) to the corresponding mean production if the firm utilized its levels of inputs most efficiently, (see Battese and Coelli [1988]). We do not consider allocative efficiency of firms in this article. Allocative and economic efficiencies have been investigated in a number of papers, including Schmidt and Lovell [1979, 1980], Kalirajan [1985], Kumbhakar [1988], Kumbhakar, Biswas, and Bailey [1989] and Bailey, et al. [1989]. We define a stochastic frontier production function model for panel data, in which technical efficiencies of firms may vary over time.

2. Time-varying model for unbalanced panel data

We consider a stochastic frontier production function with a simple exponential specification of time-varying firm effects which incorporates unbalanced panel data associated with observations on a sample of N firms over T time periods. The model is defined by

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

and

$$U_{it} = \eta_{it}\, U_i = \{\exp[-\eta(t - T)]\}U_i, \qquad t \in \mathcal{J}(i);\ i = 1, 2, \ldots, N; \tag{2}$$

where Y_{it} represents the production for the ith firm at the tth period of observation;
$f(x_{it}; \beta)$ is a suitable function of a vector, x_{it}, of factor inputs (and firm-specific variables), associated with the production of the ith firm in the tth period of observation, and a vector, β, of unknown parameters;
the V_{it}'s are assumed to be independent and identically distributed $N(0, \sigma_V^2)$ random errors;
the U_i's are assumed to be independent and identically distributed non-negative truncations of the $N(\mu, \sigma^2)$ distribution;
η is an unknown scalar parameter;
and $\mathcal{J}(i)$ represents the set of T_i time periods among the T periods involved for which observations for the ith firm are obtained.[1]

This model is such that the non-negative firm effects, U_{it}, decrease, remain constant or increase as t increases, if $\eta > 0$, $\eta = 0$ or $\eta < 0$, respectively. The case in which η is positive is likely to be appropriate when firms tend to improve their level of technical efficiency over time. Further, if the Tth time period is observed for the ith firm then $U_{iT} = U_i$, $i = 1, 2, \ldots, N$. Thus the parameters, μ and σ^2, define the statistical properties of the firm effects associated with the last period for which observations are obtained. The model assumed for the firm effects, U_i, was originally proposed by Stevenson [1980] and is a generalization of the half-normal distribution which has been frequently applied in empirical studies.

The exponential specification of the behavior of the firm effects over time (equation (2)) is a rigid parameterization in that technical efficiency must either increase at a decreasing rate ($\eta > 0$), decrease at an increasing rate ($\eta < 0$) or remain constant ($\eta = 0$). In order to permmit greater flexibility in the nature of technical efficiency, a two-parameter specification would be required. An alternative two-parameter specification, which is being investigated, is defined by

$$\eta_{it} = 1 + \eta_1(t - T) + \eta_2(t - T)^2,$$

where η_1 and η_2 are unknown parameters. This model permits firm effects to be convex or concave, but the time-invariant model is the special case in which $\eta_1 = \eta_2 = 0$.

Alternative time-varying models for firm effects have been proposed by Cornwell, Schmidt, and Sickles [1990] and Kumbhakar [1990]. Cornwell, Schmidt, and Sickles [1990] assumed that the firm effects were a quadratic function of time, in which the coefficients varied over firms according to the specifications of a multivariate distribution. Kumbhakar [1990] assumed that the non-negative firm effects, U_{it}, were the product of deterministic function of time, $\gamma(t)$, and non-negative time-invariant firm effects, U_i. The time function, $\gamma(t)$, was assumed to be defined by,

$$\gamma(t) = [1 + \exp(bt + ct^2)]^{-1}, \qquad t = 1, 2, \ldots, T.$$

This model has values for $\gamma(t)$ between zero and one and could be monotone decreasing (or increasing) or convex (or concave) depending on the values of the parameters, b and c. Kumbhakar [1990] noted that, if $b + ct$ was negative (or positive), the simpler function, $\gamma(t) = (1 + e^{bt})^{-1}$, may be appropriate.[2] The more general model of Kumbhakar [1990] would be considerably more difficult to estimate than that of the simpler exponential model of equation (2).

Given the model (1)–(2), it can be shown [see the Appendix] that the minimum-mean-squared-error predictor of the technical efficiency of the ith firm at the tth time period, $TE_{it} = \exp(-U_{it})$ is

$$E[\exp(-U_{it})|E_i] = \left\{\frac{1 - \Phi[\eta_{it}\sigma_i^* - (\mu_i^*/\sigma_i^*)]}{1 - \Phi(-\mu_i^*/\sigma_i^*)}\right\}\exp\left[-\eta_{it}\mu_i^* + \frac{1}{2}\eta_{it}^2\sigma_i^{*2}\right] \tag{3}$$

where E_i represents the $(T_i \times 1)$ vector of E_{it}'s associated with the time periods observed for the ith firm, where $E_{it} \equiv V_{it} - U_{it}$;

$$\mu_i^* = \frac{\mu\sigma_V^2 - \eta_i'E_i\sigma^2}{\sigma_V^2 + \eta_i'\eta_i\sigma^2} \tag{4}$$

$$\sigma_i^{*2} = \frac{\sigma_V^2\sigma^2}{\sigma_V^2 + \eta_i'\eta_i\sigma^2} \tag{5}$$

where η_i represents the $(T_i \times 1)$ vector of η_{it}'s associated with the time periods observed for the ith firm; and

$\Phi(\cdot)$ represents the distribution function for the standard normal random variable.

If the stochastic frontier production function (1) is of Cobb–Douglas or transcendental logarithmic type, then E_{it} is a linear function of the vector, β.

The result of equation (3) yields the special cases given in the literature. Although Jondrow, Lovell, Materov, and Schmidt [1982] only derived $E[U_i|V_i - U_i]$, the more appropriate result for cross-sectional data, $E[\exp(-U_i)|V_i - U_i]$, is obtained from equations (3)–(5) by substituting $\eta_{it} = 1 = \eta_i$ and $\mu = 0$. The special cases given in Battese and Coelli [1988] and Battese, Coelli, and Colby [1989] are obtained by substituting $\eta_i'\eta_i = T$ and $\eta_i'\eta_i = T_i$, respectively, where $\eta_{it} = 1$ (i.e., $\eta = 0$) in both cases.

Kumbhakar [1990] derived the conditional expectation of U_i, given the value of the random variables, $E_{it} \equiv V_{it} - \gamma(t)U_i$, $t = 1, 2, \ldots, T$, under the assumptions that the U_i's had half-normal distribution. Kumbhakar's [1990] model also accounted for the presence of allocative inefficiency, but gave no empirical application.

The mean technical efficiency of firms at the tth time period,

$$TE_t \equiv E[\exp(-\eta_t U_i)], \qquad \text{where } \eta_t = \exp[-\eta(t - T)],$$

obtained by straightforward integration with the density function of U_i, is

$$TE_t = \left\{\frac{1 - \Phi[\eta_t\sigma - (\mu/\sigma)]}{[1 - \Phi(-\mu/\sigma)]}\right\}\exp\left[-\eta_t\mu + \frac{1}{2}\,\eta_t^2\sigma^2\right]. \tag{6}$$

If the firm effects are time invariant, then the mean technical efficiency of firms in the industry is obtained from equation (6) by substitution of $\eta_t = 1$. This gives the result presented in equation (8) of Battese and Coelli [1988].

Operational predictors for equations (3) and (6) may be obtained by substituting the relevant parameters by their maximum-likelihood estimators. The maximum-likelihood estimates for the parameters of the model and the predictors for the technical efficiencies of firms can be approximated by the use of the computer program, FRONTIER, which was written by Tim Coelli.[3] The likelihood function for the sample observations, given the parameterization of the model (1)–(2) used in FRONTIER, is presented in the Appendix.

3. Empirical example

Battese, Coelli, and Colby [1989] used a set of panel data on 38 farmers from an Indian village to estimate the parameters of a stochastic frontier production function for which the technical efficiencies of individual farmers were assumed to be time invariant. We consider a subset of these data for those farmers, who had access to irrigation and grew paddy, to estimate a stochastic frontier production frontier with time-varying firm effects, as specified by equations (1)–(2) in Section 2. The data were collected by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) from farmers in the village of Aurepalle. We consider the data for fifteen farmers who engaged in growing paddy for between four and ten years during the period, 1975–1976 through 1984–1985. Nine of the fifteen farmers were observed for all the ten years involved. A total of 129 observations were used, so 21 observations were missing from the panel.

The stochastic frontier production function for the panel data on the paddy farmers in Aurepalle which we estimate is defined by

$$\log(Y_{it}) = \beta_0 + \beta_1 \log(Land_{it}) + \beta_2(IL_{it}/Land_{it}) + \beta_3 \log(Labor_{it}) + \beta_4 \log(Bullock_{it}) + \beta_5 \log(Costs_{it}) + V_{it} - U_{it}, \quad (7)$$

where the subscripts i and t refer to the *i*th farmer and the *t*th observation, respectively;

Y represents the total value of output (in Rupees) from paddy and any other crops which might be grown;

Land represents the total area (in hectares) of irrigated and unirrigated land, denoted by IL_{it} and UL_{it}, respectively;

Labor represents the total number of hours of human labor (in male equivalent units)[4] for family members and hired laborers;

Bullock represents the total number of hours of bullock labor for owned or hired bullocks (in pairs);

Costs represents the total value of input costs involved (fertilizer, manure, pesticides, machinery, etc.); and

V_{it} and U_{it} are the random variables whose distributional properties are defined in Section 2.

A summary of the data on the different variables in the frontier production function is given in table 1. It is noted that about 30 percent of the total land operated by the paddy farmers in Aurepalle was irrigated. Thus the farmers involved were generally also engaged in dryland farming. The minimum value of irrigated land was zero because not all the farmers involved grew paddy (irrigated rice) in all the years involved.

The production function, defined by equation (7), is related to the function which was estimated in Battese, Coelli, and Colby [1989, p. 333], but family and hired labor are aggregated (i.e., added).[5] The justification for the functional form considered in Battese, Coelli, and Colby [1989] is based on the work of Bardhan [1973] and Deolalikar and Vijverberg

Table 1. Summary statistics for variables in the stochastic frontier production function for paddy farmers in Aurepalle.[1]

Variable	Sample Mean	Sample Standard Deviation	Minimum Value	Maximum Value
Value of Output (Rupees)	6939	4802	36	18094
Total Land (hectares)	6.70	4.24	0.30	20.97
Irrigated Land (hectares)	1.99	1.47	0.00	7.09
Human Labor (hours)	4126	2947	92	6205
Bullock Labor (hours)	900.4	678.2	56	4316
Input Costs (Rupees)	1273	1131	0.7	6205

[1]The data, consisting of 129 observations for each variable, collected from 15 paddy farmers in Aurepalle over the ten-year period, 1975–1976 to 1984–1985, were collected by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) as part of its Village Level Studies (see Binswanger and Jodha [1978]).

[1983] with Indian data on hired and family labor and irrigated and unirrigated land. The production function of equation (7) is a linearized version of that which was directly estimated in Battese, Coelli, and Colby [1989][6] (see the model in Defourny, Lovell, and N'gbo [1990]).

The original values of output and input costs used in Batese, Coelli, and Colby [1989] are deflated by a price index for the analyses in this article. The price index used was constructed using data, supplied by ICRISAT, on prices and quantities of crops grown in Aurepalle.

The stochastic frontier model, defined by equation (7), contains six β-parameters and the four additional parameters associated with the distributions of the V_{it}- and U_{it}-random variables. Maximum-likelihood estimates for these parameters were obtained by using the computer program, FRONTIER. The frontier function (7) is estimated for five basic models:

Model 1.0 involves all parameters being estimated;
Model 1.1 assumes that $\mu = 0$;
Model 1.2 assumes that $\eta = 0$;
Model 1.3 assumes that $\mu = \eta = 0$; and
Model 1.4 assumes that $\gamma = \mu = \eta = 0$.

Model 1.0 is the stochastic frontier production function (7) in which the farm effects, U_{it}, have the time-varying structure defined in Section 2 (i.e., η is an unknown parameter and the U_i's of equation (2) are non-negative truncations of the $N(\mu, \sigma^2)$ distribution). Model 1.1 is the special case of Model 1.0 in which the U_i's have half-normal distribution (i.e., μ is assumed to be zero). Model 1.2 is the time-invariant model considered by Battese, Coelli, and Colby [1989]. Model 1.3 is the time-invariant model in which the farm effects, U_i, have half-normal distribution. Finally, Model 1.4 is the traditional average response function in which farms are assumed to be fully technically efficient (i.e., the farm effects, U_{it}, are absent from the model).

Empirical results for these five models are presented in table 2. Tests of hypotheses involving the parameters of the distributions of the U_{it}-random variables (farm effects) are obtained by using the generalized likelihood-ratio statistic. Several hypotheses are considered for different distributional assumptions and the relevant statistics are presented in table 3.

Given the specifications of the stochastic frontier with time-varying farm effects (Model 1.0), it is evident that the traditional average production function is not an adequate representation of the data (i.e., the null hypothesis, H_0: $\gamma = \mu = \eta = 0$, is rejected). Further, the hypotheses that time-invariant models for farm effects apply are also rejected (i.e., both H_0: $\mu = \eta = 0$ and H_0: $\eta = 0$ would be rejected). However, the hypothesis that the half-normal distribution is an adequate representation for the distribution of the farm effects is not rejected using these data. Given that the half-normal distribution is assumed appropriate to define the distribution of the farm effects, the hypothesis that the yearly farm effects are time invariant is also rejected by the data.

On the basis of these results it is evident that the hypothesis of time-invariant technical efficiencies of paddy farmers in Aurepalle would be rejected. Given the specifications of Model 1.1 (involving the half-normal distribution), the technical efficiencies of the individual paddy farmers are calculated using the predictor, defined by equation (3). The values obtained, together with the estimated mean technical efficiencies (obtained using equation (6)) in the ten years involved, are presented in table 4.

Table 2. Maximum-likelihood estimates for parameters of stochastic frontier production functions for Aurepalle paddy farmers.[1]

		MLE Estimates for Models				
Variable	Parameter	Model 1.0	Model 1.1	Model 1.2	Model 1.3	Model 1.4
Constant	β_0	3.74 (0.96)	3.86 (0.94)	3.90 (0.73)	3.87 (0.68)	3.71 (0.66)
log(Land)	β_1	0.61 (0.23)	0.63 (0.20)	0.63 (0.15)	0.63 (0.15)	0.62 (0.15)
IL/Land	β_2	0.81 (0.43)	1.05 (0.33)	0.90 (0.30)	0.89 (0.29)	0.80 (0.27)
log(Labor)	β_3	0.76 (0.21)	0.74 (0.18)	0.74 (0.15)	0.74 (0.14)	0.74 (0.14)
log(Bullocks)	β_4	−0.45 (0.16)	−0.43 (0.11)	−0.44 (0.11)	−0.44 (0.11)	−0.43 (0.12)
log(Costs)	β_5	0.079 (0.048)	0.058 (0.038)	0.052 (0.042)	0.052 (0.042)	0.053 (0.043)
	$\sigma_S^2 \equiv \sigma_V^2 + \sigma^2$	0.129 (0.048)	0.104 (0.010)	0.136 (0.040)	0.142 (0.028)	0.135 (0.019)
	$\gamma \equiv \sigma^2/\sigma_S^2$	0.22 (0.21)	0.056 (0.012)	0.11 (0.26)	0.14 (0.17)	0
	μ	−0.77 (1.79)	0	−0.07 (0.43)	0	0
	η	0.27 (0.97)	0.138 (0.047)	0	0	0
	Log (likelihood)	−40.788	−40.798	−50.408	−50.416	−50.806

[1]The estimated standard errors for the parameter estimators are presented below the corresponding estimates. These values are generated by the computer program, FRONTIER.

Table 3. Tests of hypotheses for parameters of the distribution of the farm effects, U_{it}.

Assumptions	Null Hypothesis H_0	χ^2-statistic	$\chi^2_{0.95}$-value	Decision
Model 1.0	$\gamma = \mu = \eta = 0$	20.04	7.81	Reject H_0
Model 1.0	$\mu = \eta = 0$	19.26	5.99	Reject H_0
Model 1.0	$\mu = 0$	0.02	3.84	Accept H_0
Model 1.0	$\eta = 0$	19.24	3.84	Reject H_0
Model 1.1 ($\mu = 0$)	$\gamma = \eta = 0$	20.02	5.99	Reject H_0
Model 1.1 ($\mu = 0$)	$\eta = 0$	19.24	3.84	Reject H_0

Table 4. Predicted technical efficiencies of paddy farmers in Aurepalle for the years 1975–1976 through 1984–1985.[1]

Farmer Number	75–76	76–77	77–78	78–79	79–80	80–81	81–82	82–83	83–84	84–85
1	.861	.878	.892	.905	.916	.927	.936	.944	.951	.957
2	.841	.859	.876	.891	.904	.915	.926	.935	.943	.950
3	.569	.611	.651	.687	.721	.752	—	—	—	—
4	.549	.593	.633	.671	.706	.738	.767	.794	.818	.839
5	.711	.743	.771	.797	.820	.841	.860	.876	.891	.904
6	.798	.821	.842	.860	.877	.891	.905	.916	.926	.935
7	.576	.618	.657	.693	.726	.756	.784	.808	.831	—
8	.776	.801	.823	—	.862	.878	.893	.906	.917	.927
9	.575	.617	.656	.692	.725	.756	.783	.808	.830	.850
10	.862	.878	.892	.905	.917	.927	.936	.944	.951	.957
11	.778	.803	.825	.846	.864	.880	.894	.907	.918	.928
12	.712	.743	.771	.797	.820	.841	.860	.876	.891	.904
13	.601	.678	.712	.743	.772	.798	.821	—	—	—
14	—	.789	.813	.834	.853	—	—	—	—	—
15	—	—	—	—	—	—	.908	.919	.929	.938
Mean	.821	.841	.859	.875	.890	.903	.915	.925	.934	.942

[1]In years when particular farmers were not observed, no values of technical efficiencies are calculated.

The technical efficiencies range between 0.549 and 0.862 in 1975–1976 and, between 0.839 and 0.957 in 1984–1985. Because the estimate for the parameter, η, is positive ($\hat{\eta} = 0.138$) the technical efficiencies increase over time, according to the assumed exponential model, defined by equation (2). These predicted technical efficiencies of the 15 paddy farmers are graphed against year of observation in figure 1. These data indicate that there exist considerable variation in the efficiencies of the paddy farmers, particularly at the beginning of the sample period. Given the assumption that the farm effects change exponentially over time, it is expected that the predicted efficiencies converge over a period of generally increasing levels of technical efficiency.

The above results are, however, based on the stochastic frontier production function (7), which assumes that the parameters are time invariant. In particular, the presence of technical progress is not accounted for in the model. Given that year of observation is included as an additional explanatory variable, then the estimated stochastic frontier production function is

$$\log Y = \underset{(1.75)}{2.80} + \underset{(0.37)}{0.50} \log(\text{Land}) + \underset{(0.47)}{0.53} (\text{IL/Land}) + \underset{(0.32)}{0..91} \log(\text{Labor})$$

$$- \underset{(0.098)}{0.489} \log(\text{Bullocks}) + \underset{(0.040)}{0.051} \log(\text{Costs}) + \underset{(0.019)}{0.050} \text{ Year} \tag{8}$$

where $\hat{\sigma}_S^2 = \underset{(0.084)}{0.130}, \quad \hat{\gamma} = \underset{(0.44)}{0.21}, \quad \hat{\mu} = \underset{(0.98)}{-0.69}, \quad \hat{\eta} = \underset{(0.65)}{0.11}$

and log (likelihood) = − 38.504.

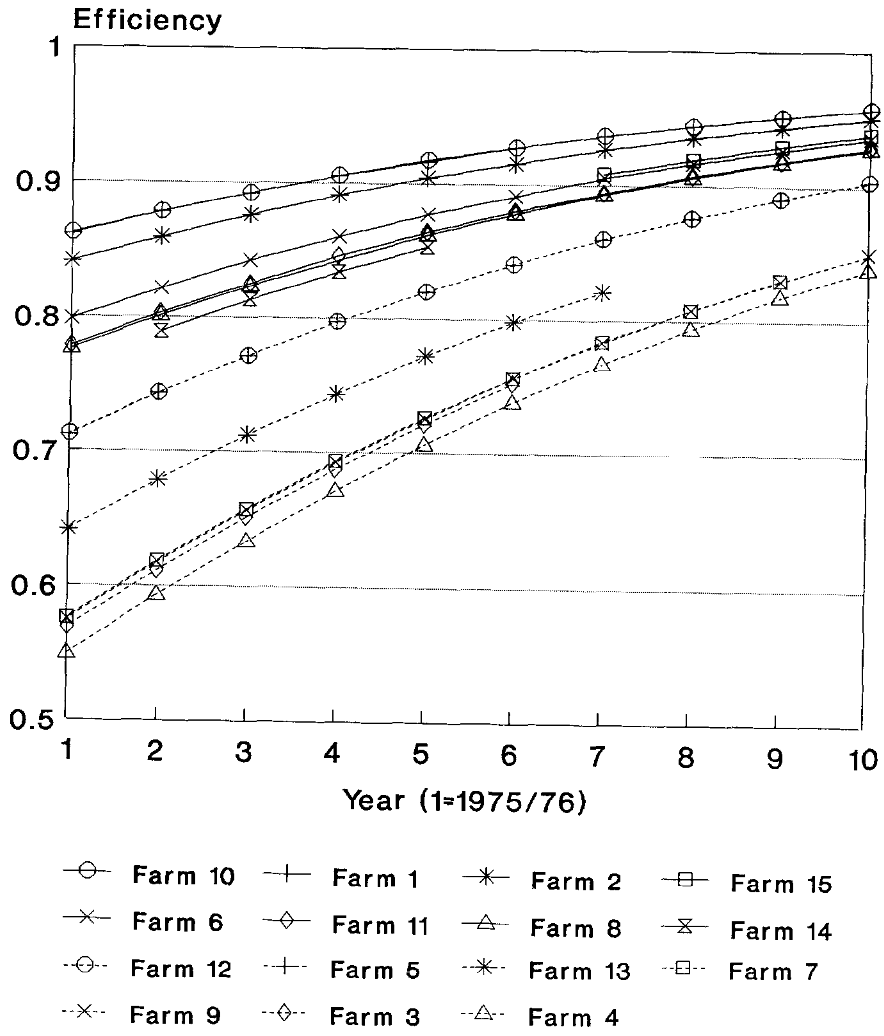

Figure 1. Predicted technical efficiencies.

Generalized likelihood-ratio tests of the hypotheses that the parameters, μ, η and γ, are zero (individually or jointly) yield insignificant results. Thus the inclusion of the year of observation in the model (i.e., Hicksian neutral technological change), leads not only to the conclusion that technical efficiency of the paddy farmers is time invariant, but that the stochastic frontier production function is not significantly different from the traditional average response model. This response function is estimated by

$$\begin{aligned}\log Y = {} & \underset{(0.63)}{2.73} + \underset{(0.13)}{0.51}\log(\text{Land}) + \underset{(0.26)}{0.50}(\text{IL/Land}) + \underset{(0.14)}{0.91}\log(\text{Labor}) \\ & - \underset{(0.11)}{0.48}\log(\text{Bullocks}) + \underset{(0.040)}{0.048}\log(\text{Costs}) + \underset{(0.011)}{0.054}\,\text{Year}\end{aligned} \tag{9}$$

where $\hat{\sigma}_V^2 = 0.113$ and log (likelihood) $= -38.719$.

The estimated response function in equation (9) is such that the returns-to-scale parameter is estimated by 0.990 which is not significantly different from one, because the estimated standard error of the estimator is 0.065. Thus, the hypothesis of constant returns to scale for the paddy farmers would not be rejected using these data.

The coefficient of the ratio of irrigated land to total land operated, IL/Land, is significantly different from zero. Using the estimates for the elasticity of land and the coefficient of the land ratio, one hectare of irrigated land is estimated to be equivalent to about 1.98 hectares of unirrigated land for Aurepalle farmers who grow paddy and other crops.[7] This compares with 3.50 hectares obtained by Battese, Coelli, and Colby [1989] using data on all 38 farmers in Aurepalle. The smaller value obtained using only data on paddy farmers is probably due to the smaller number of unirrigated hectares in this study than in the earlier study involving all farmers in the village.

The estimated elasticity for bullock labor on paddy farms is negative. This result was also observed in Saini [1979] and Battese, Coelli, and Colby [1989]. A plausible argument for this result is that paddy farmers may use bullocks more in years of poor production (associated with low rainfall) for the purpose of weed control, levy bank maintenance, etc., which are difficult to conduct in years of higher rainfall and higher output. Hence, the bullock-labor variable may be acting as an inverse proxy for rainfall.

The coefficient, 0.054, of the variable, year of observation, in the estimated response function, given by equation (9), implies that value of output (in real terms) is estimated to have increased by about 5.4 percent over the ten-year period for the paddy farmers in Aurepalle.

4. Conclusions

The empirical application of the stochastic frontier production function model with time-varying firm effects (1)–(2), in the analysis of data from paddy farmers in an Indian village, revealed that the technical efficiencies of the farmers were not time invariant when year of observation was excluded from the stochastic frontier. However, the inclusion of year of observation in the frontier model led to the finding that the corresponding technical efficiencies were time invariant. In addition, the stochastic frontier was not significantly different from the traditional average response function. This implies that, given the state of technology among paddy farmers in the Indian village involved, technical inefficiency is not an issue of significance provided technical change is accounted for in the empirical analysis. However, in other empirical applications of the time-varying model which we have conducted (see Battese and Tessema [1992]), the inclusion of time-varying parameters in the stochastic frontier has not necessarily resulted in time-invariant technical efficiencies or the conclusion that technical inefficiency does not exist.

The stochastic frontier production function estimated in Section 3 did not involve farmer-specific variables. To the extent that farmer- (and farm-) specific variables influence technical efficiencies, the empirical analysis presented in Section 3 does not appropriately predict technical efficiencies. More detailed modeling of the variables influencing production and the statistical distribution of the random variables involved will lead to improved analysis of production and better policy decisions concerning productive activity. We are confident that further theoretical developments in stochastic frontier modeling and the prediction of technical efficiencies of firms will assist such practical decision making.

Appendix

Consider the frontier production function[8]

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

where

$$E_{it} = V_{it} - \eta_{it}U_i \tag{A.2}$$

and

$$\eta_{it} = e^{-\eta(t-T)}, \qquad t \in \mathcal{J}(i);\ i = 1, 2, \ldots, N. \tag{A.3}$$

It is assumed that the V_{it}'s are iid $N(0, \sigma_V^2)$ random variables, independent of the U_i's, which are assumed to be non-negative truncations of the $N(\mu, \sigma^2)$ distribution.

The density function for U_i is

$$f_{U_i}(u_i) = \frac{\exp\left[-\frac{1}{2}(u_i - \mu)^2/\sigma^2\right]}{(2\pi)^{1/2}\sigma[1 - \phi(-\mu/\sigma)]}, \qquad u_i \geq 0, \tag{A.4}$$

where $\Phi(\cdot)$ represents the distribution function for the standard normal random variable.

It can be shown that the mean and variance of U_i are[9]

$$E(U_i) = \mu + \sigma\{\phi(-\mu/\sigma)/[1 - \Phi(-\mu/\sigma)]\} \tag{A.5}$$

and

$$\mathrm{Var}(U_i) = \sigma^2\left\{1 - \frac{\phi(-\mu/\sigma)}{1 - \Phi(-\mu/\sigma)}\left[\frac{\mu}{\sigma} + \frac{\phi(-\mu/\sigma)}{1 - \Phi(-\mu/\sigma)}\right]\right\}, \tag{A.6}$$

where $\phi(\cdot)$ represents the density function for the standard normal distribution.

From the joint density function for U_i and V_i, where V_i represents the $(T_i \times 1)$ vector of the V_{it}'s associated with the T_i observations for the *i*th firm, it follows readily that the joint density function for U_i and E_i, where E_i is the $(T_i \times 1)$ vector of the values of $E_{it} \equiv V_{it} - \eta_{it}U_i$, is

$$f_{U_i,E_i}(u_i, e_i) = \frac{\exp - \frac{1}{2}\{[(u_i - \mu)^2/\sigma^2] + [(e_i + \eta_i u_i)'(e_i + \eta_i u_i)/\sigma_V^2]\}}{(2\pi)^{(T_i+1)/2}\sigma\,\sigma_V^{T_i}[1 - \Phi(-\mu/\sigma)]} \tag{A.7}$$

where e_i is a possible value for the random vector, E_i.

The density function for E_i, obtained by integrating $f_{U_i,E_i}(u_i, e_i)$ with respect to the range for U_i, namely $u_i \geq 0$, is

$$f_{E_i}(e_i) = \frac{[1 - \Phi(-\mu_i^*/\sigma_i^*)]\exp - \frac{1}{2}\{(e_i'e_i/\sigma_V^2) + (\mu/\sigma)^2 - (\mu_i^*/\sigma_i^*)^2\}}{(2\pi)^{T_i/2}\sigma_V^{(T_i-1)}[\sigma_V^2 + \eta_i'\eta_i\sigma^2]^{1/2}[1 - \Phi(-\mu/\sigma)]} \tag{A.8}$$

where

$$\mu_i^* \equiv \frac{\mu\sigma_V^2 - \eta_i'e_i\sigma^2}{\sigma_V^2 + \eta_i'\eta_i\sigma^2} \tag{A.9}$$

and

$$\sigma_i^{*2} \equiv \frac{\sigma^2\sigma_V^2}{\sigma_V^2 + \eta_i'\eta_i\sigma^2}. \tag{A.10}$$

From the above results, it follows that the conditional density function of U_i, given that the random vector, E_i, has value, e_i, is

$$f_{U_i|E_i=e_i}(u_i) = \frac{\exp - \frac{1}{2}[(u_i - \mu_i^*)/\sigma_i^*]^2}{(2\pi)^{1/2}\sigma_i^*[1 - \Phi(-\mu_i^*/\sigma_i^*)]},\ u_i \geq 0. \tag{A.11}$$

This is the density function of the positive truncation of the $N(\mu_i^*, \sigma_i^{*2})$ distribution. Since the conditional expectation of $\exp(-\eta_{it}U_i)$, given $E_i = e_i$, is defined by

$$E\{\exp(-\eta_{it}U_i|E_i = e_i)\} = \int_0^\infty \exp(-\eta_{it}u_i)f_{U_i|E_i=e_i}(u_i)du_i,$$

the result of equation (3) of the text of this article is obtained by straightforward integral calculus.

If the frontier production function (A.1)-(A.3) is appropriate for production, expressed in the original units of output, then the prediction of the technical efficiency of the *i*th firm at the time of the tth observation, $TE_{it} = 1 - (\eta_{it} U_i / x_{it}\beta)$, requires the conditional expectation of U_i, given $E_i = e_i$. This can be shown to be

$$E(U_i|E_i = e_i) = \mu_i^* + \sigma_i^* \{\phi(-\mu_i^*/\sigma_i^*)/[1 - \Phi(-\mu_i^*/\sigma_i^*)]\} \tag{A.12}$$

where μ_i^* and σ_i^{*2} are defined by equations (A.9) and (A.10), respectively.

The density function for Y_i, the $(T_i \times 1)$ random vector of Y_{it}'s for the ith firm, is obtained from (A.8) by substituting $(y_i - x_i\beta)$ for e_i, where x_i is the $(T_i \times k)$ matrix of x_{it}'s for the ith firm, where k is the dimension of the vector, β. The logarithm of the likelihood function for the sample observations, $y \equiv (y_1', y_2', \ldots, y_N')'$, is thus

$$\begin{aligned} L^*(\theta^*; y) = & -\frac{1}{2}\left(\sum_{i=1}^{N} T_i\right) \ell n(2\pi) - \frac{1}{2}\sum_{i=1}^{N}(T_i - 1)\ell n(\sigma_V^2) - \frac{1}{2}\sum_{i=1}^{N}\ell n(\sigma_V^2 + \eta_i'\eta_i\sigma^2) \\ & - N\,\ell n[1 - \Phi(-\mu/\sigma)] + \sum_{i=1}^{N}\ell n[1 - \Phi(-\mu_i^*/\sigma_i^*)] \\ & - \frac{1}{2}\sum_{i=1}^{N}[(y_i - x_i\beta)'(y_i - x_i\beta)/\sigma_V^2] - \frac{1}{2}N(\mu/\sigma)^2 + \frac{1}{2}\sum_{i=1}^{N}(\mu_i^*/\sigma_i^*)^2, \end{aligned} \tag{A.13}$$

where $\theta^* \equiv (\beta', \sigma_V^2, \sigma^2, \mu, \eta)'$.

Using the reparameterization of the model, suggested by Battese and Corra [1977], where $\sigma_V^2 + \sigma^2 = \sigma_S^2$ and $\gamma = \sigma^2/\sigma_S^2$, the logarithm of the likelihood function is expressed by

$$\begin{aligned} L^*(\theta; y) = & -\frac{1}{2}\left(\sum_{i=1}^{N} T_i\right)\{\ell n(2\pi) + \ell n(\sigma_S^2)\} - \frac{1}{2}\sum_{i=1}^{N}(T_i - 1)\ell n(1 - \gamma) \\ & - \frac{1}{2}\sum_{i=1}^{N}\ell n[1 + (\eta_i'\eta_i - 1)\gamma] - N\ell n[1 - \Phi(-z)] - \frac{1}{2}Nz^2 \\ & + \sum_{i=1}^{N}\ell n[1 - \Phi(-z_i^*)] + \frac{1}{2}\sum_{i=1}^{N} z_i^{*2} \\ & - \frac{1}{2}\sum_{i=1}^{N}(y_i - x_i\beta)'(y_i - x_i\beta)/(1 - \gamma)\sigma_S^2, \end{aligned} \tag{A.14}$$

where $\theta \equiv (\beta', \sigma_S^2, \gamma, \mu, \eta)'$, $z \equiv \mu/(\gamma\sigma_S^2)^{1/2}$ and

$$z_i^* = \frac{\mu(1 - \gamma) - \gamma\eta_i'(y_i - x_i\beta)}{\{\gamma(1 - \gamma)\sigma_S^2[1 + (\eta_i'\eta_i - 1)\gamma]\}^{1/2}}.$$

The partial derivations of the loglikelihood function (A.14) with respect to the parameters, β, σ_S^2, γ, μ and η, are given by

$$\frac{\partial L^*}{\partial \beta} = \sum_{i=1}^{N} x'_i(y_i - x_i\beta)[(1-\gamma)\sigma_S^2]^{-1}$$

$$+\sum_{i=1}^{N}\left[\frac{\phi(-z_i^*)}{1-\Phi(-z_i^*)} + z_i^*\right]\gamma\, x_i'\eta_i\{\gamma(1-\gamma)\sigma_S^2[1 + (\eta_i'\eta_i - 1)\gamma]\}^{-1/2}$$

$$\frac{\partial L^*}{\partial \sigma_S^2} = -\frac{1}{2\sigma_S^2}\left\{\sum_{i=1}^{N} T_i - N\left[\frac{\phi(-z)}{1-\Phi(-z)} + z_3\right]z + \sum_{i=1}^{N}\left[\frac{\phi(-z_i^*)}{1-\Phi(-z_i^*)} + z_i^*\right]z_i^*\right.$$

$$\left. - \sum_{i=1}^{N}(y_i - x_i\beta)'(y_i - x_i\beta)[(1-\gamma)\sigma_S^2]^{-1}\right\}$$

$$\frac{\partial L^*}{\partial \gamma} = \frac{(1-\gamma)^{-1}}{2}\sum_{i=1}^{N}(T_i - 1) - \frac{1}{2}\sum_{i=1}^{N}(\eta_i'\eta_i - 1)[1 + (\eta_i'\eta_i - 1)\gamma]^{-1}$$

$$+\frac{N}{2}\left[\frac{\phi(-z)}{1-\Phi(-z)} + z\right]z\gamma^{-1} + \sum_{i=1}^{N}\left[\frac{\phi(-z_i^*)}{1-\Phi(-z_i^*)} + z_i^*\right]\frac{\partial z_i^*}{\partial \gamma}$$

$$-\frac{1}{2}\sum_{i=1}^{N}(y_i - x_i\beta)'(y_i - x_i\beta)[(1-\gamma)\sigma_S]^{-2}$$

$$\frac{\partial L^*}{\partial \mu} = -\frac{N}{(\gamma\sigma_S^2)^{1/2}}\left[\frac{\phi(-z)}{1-\Phi(-z)} + z\right] + \sum_{i=1}^{N}\left[\frac{\phi(-z_i^*)}{1-\Phi(-z_i^*)} + z_i^*\right]$$

$$\times\frac{(1-\gamma)}{\{\gamma(1-\gamma)\sigma_S^2[1 + (\eta_i'\eta_i - 1)\gamma\}^{1/2}}$$

$$\frac{\partial L^*}{\partial \eta} = \sum_{i=1}^{N}\left[\frac{\phi(-z_i^*)}{1-\Phi(-z_i^*)} + z_i^*\right]\frac{\partial z_i^*}{\partial \eta} - \frac{\gamma}{2}\sum_{i=1}^{N}\frac{\partial \eta_i'\eta_i}{\partial \eta}[1 + (\eta_i'\eta_i - 1)\gamma]^{-1}$$

where

$$\frac{\partial z_i^*}{\partial \gamma} = -\frac{[\mu + \eta_i'(y_i - x_i\beta)]}{\sigma_S\{\gamma(1-\gamma)[1 + (\eta_i'\eta_i - 1)\gamma]\}^{1/2}}$$

$$-\frac{1}{2}\frac{[\mu(1-\gamma) - \gamma\eta_i'(y_i - x_i\beta)][(1-2\gamma) + (\eta_i'\eta_i - 1)\gamma(2-3\gamma)]}{\sigma_S\{\gamma(1-\gamma)[1 + (\eta_i'\eta_i - 1)\gamma]\}^{3/2}}$$

$$\frac{\partial z_i^*}{\partial \eta} = \frac{\gamma \sum_{t \in \mathcal{I}(i)} (t - T)e^{-\eta(t-T)}(y_{it} - x_{it}\beta)}{\{\gamma(1 - \gamma)\sigma_S^2[1 + (\eta_i'\eta_i - 1)\gamma]\}^{1/2}}$$

$$- \frac{[\mu(1 - \gamma) - \gamma\eta_i'(y_i - x_i\beta)] \frac{1}{2} \gamma^2(1 - \gamma)\sigma_S^2 \frac{\partial \eta_i'\eta_i}{\partial \eta}}{\{\gamma(1 - \gamma)\sigma_S^2[1 + (\eta_i'\eta_i - 1)\gamma]\}^{3/2}}$$

and

$$\frac{\partial \eta_i'\eta_i}{\partial \eta} = -2 \sum_{t \in \mathcal{I}(i)} (t - T)e^{-2\eta(t-T)} \qquad \text{if } \eta \neq 0.$$

Notes

1. If the ith firm is observed in all the T time periods involved, then $\mathcal{I}(i) = \{1, 2, \ldots, T\}$. However, if the ith firm was continuously involved in production, but observations were only obtained at discrete intervals, then $\mathcal{I}(i)$ would consist of a subset of the integers, 1, 2, . . ., T, representing the periods of observations involved.
2. It is somewhat unusual that the value of $\gamma(t)$ for the period before the first observation, $t = 0$, is 0.5.
3. The original version of FRONTIER (see Coelli [1989]) was written to estimate the time-invariant panel data model presented in Battese and Coelli [1988]. It was amended to account for unbalanced panel data and applied in Battese, Coelli, and Colby [1989]. Recently, FRONTIER was updated to estimate the time-varying model defined by equations (1) and (2), (see Coelli [1991, 1992]). FRONTIER Version 2.0 is written in Fortran 77 for use on IBM compatible PC's. The source code and executable program are available from Tim Coelli on a 5.25 inch disk.
4. Labor hours were converted to male equivalent units according to the rule that female and child hours were considered equivalent to 0.75 and 0.50 male hours, respectively. These ratios were obtained from ICRISAT.
5. The hypothesis that family and hired labor were equally productive was tested and accepted in Battese, Coelli, and Colby [1989]. Hence only total labor hours are considered in this paper.
6. The deterministic component of the stochastic frontier production function estimated in Battese, Coelli, and Colby [1989], considering only the land variable (consisting of a weighted average of unirrigated land and irrigated land), is defined by,

   $$Y = a_0[a_1 UL + (1 - a_1)IL]^{\beta_1}.$$

   This model is expressed in terms of Land $\equiv$ UL + IL and IL/Land, as follows

   $$Y = a_0 \times a_1^{\beta_1} (\text{Land})^{\beta_1}[1 + (b_1 - 1)(\text{IL/Land})]^{\beta_1}, \text{ where } b_1 = (1 - a_1)/a_1.$$

   By taking logarithms of both sides and considering only the first term of the infinite series expansions of the function involving the land ratio, IL/Land, we obtain

   $$\log Y \doteq \text{constant} + \beta_1 \log(\text{Land}) + \beta_2 (\text{IL/Land}), \text{ where } \beta_2 = \beta_1(b_1 - 1).$$

7. The calculations involved are: $\hat{\beta}_1 = 0.512$, $\hat{\beta}_2 \equiv \hat{\beta}_1(\hat{b}_1 - 1) = 0.501$ implies $\hat{b}_1 = 1.98$, where b_1 is the value of one hectare of irrigated land in terms of unirrigated land for farmers who grow paddy and other crops.
8. In the frontier model (2), the notation, Y_{it}, represented the actual production at the time of the tth observation for the ith firm. However, given that (2) involves a Cobb–Douglas or transcendental logarithmic model, then Y_{it} and x_{it} in this Appendix would represent logarithms of output and input values, respectively.

9. We prefer not to use the notation, σ_U^2, for the variance of the normal distribution which is truncated (at zero) to obtain the distribution of the non-negative firm effects, because this variance is *not* the variance of U_i. For the case of the half-normal distribution the variance of U_i is $\sigma^2(\pi - 2)/\pi$. This fact needs to be kept in mind in the interpretation of empirical results for the stochastic frontier model.

References

Afriat, S.N. (1972). "Efficiency Estimation of Production Functions." *International Economic Review* 13, pp. 568–598.

Aigner, D.J. and S.F. Chu. (1968). "On Estimating the Industry Production Function." *American Economic Review* 58, pp. 826–839.

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

Bailey, D.V., B. Biswas, S.C. Kumbhakar, and B.K. Schulthies. (1989). "An Analysis of Technical, Allocative, and Scale Inefficiency: The Case of Ecuadorian Dairy Farms." *Western Journal of Agricultural Economics* 14, pp. 30–37.

Bardhan, P.K. (1973). "Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture." *Journal of Political Economy* 81, pp. 1370–1386.

Battese, G.E. (1991). "Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics." *Agricultural Economics* (to appear).

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, pp. 387–399.

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, pp. 327–348.

Battese, G.E. and G.S. Corra. (1977). "Estimation of a Production Frontier Model: With Application to the Pastoral Zone of Eastern Australia." *Australian Journal of Agricultural Economics* 21, pp. 169–179.

Battese, G.E. and G.A. Tessema. (1992). "Estimation of Stochastic Frontier Production Functions with Time-Varying Parameters and Technical Efficiencies Using Panel Data from Indian Villages." Paper presented at the 36th Annual Conference of the Australian Agricultural Economics Society, Canberra, 10–12 February, 1992.

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

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

Binswanger, H.P. and N.S. Jodha. (1978). *Manual of Instructions for Economic Investigators in ICRISAT's Village Level Studies* Volume II. Village Level Studies Series, Economics Program, International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Andhra Pradesh, India.

Coelli, T.J. (1989). "Estimation of Frontier Production Functions: A Guide to the Computer Program, FRONTIER." *Working Papers in Econometrics and Applied Statistics* no. 34, Department of Econometrics, University of New England, Armidale, p. 31.

Coelli, T.J. (1991). "Maximum-Likelihood Estimation of Stochastic Frontier Production Functions with Time-Varying Technical Efficiency Using the Computer Program, FRONTIER Version 2.0." *Working Papers in Econometrics and Applied Statistics* no. 57, Department of Econometrics, University of New England, Armidale, p. 45.

Coelli, T.J. (1992). "A Computer Program for Frontier Production Function Estimation." *Economics Letters* (to appear).

Cornwell, C., P. Schmidt, and R.C. Sickles. (1990). "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics* 46, pp. 185–200.

Defourny, J., C.A.K. Lovell, and A.G.M. N'Gbo. (1990). "Variation in Productive Efficiency in French Workers' Cooperatives." *Journal of Productivity Analysis* 3(1/2), pp. 103–117.

Deolalikar, A.B. and W.P.M. Vijverberg. (1983). "The Heterogeneity of Family and Hired Labor in Agricultural Production: A Test Using District-Level Data from India." *Journal of Economic Development* 8(2), pp. 45–69.

Färe, R., S. Grosskopf, and C.A.K. Lovell. (1985). *The Measurement of Efficiency of Production*. Dordrecht: Kluwer-Nijhoff.

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, pp. 5–25.

Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt. (1982). "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *Journal of Econometrics* 19, pp. 233–238.

Kalirajan, K.P. (1985). "On Measuring Absolute Technical and Allocative Efficiencies." *Sankhya: The Indian Journal of Statistics*, Series B, 47, pp. 385–400.

Kumbhakar, S.C. (1988). "On the Estimation of Technical and Allocative Inefficiency Using Stochastic Frontier Functions: The Case of U.S. Class 1 Railroads." *International Economic Review* 29, pp. 727–743.

Kumbhakar, S.C. (1990). "Production Frontiers, Panel Data and Time-Varying Technical Inefficiency." *Journal of Econometrics* 46, pp. 201–211.

Kumbhakar, S.C., B. Biswas, and D.V. Bailey. (1989). "A Study of Economic Efficiency and Utah Dairy Farms: A System Approach." *The Review of Economics and Statistics* 71(4), pp. 595–604.

Ley, E. (1990). "A Bibliography on Production and Efficiency." mimeo, Department of Economics, University of Michigan, Ann Arbor, MI 48109, p. 32.

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

Richmond, J. (1974). "Estimating the Efficiency of Production." *International Economic Review* 15, pp. 515–521.

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

Schmidt, P. (1976). "On the Statistical Estimation of Parametric Frontier Production Functions." *The Review of Economics and Statistics* 58, pp. 238–239.

Schmidt, P. (1986). "Frontier Production Functions." *Economic Reviews* 4, pp. 289–328.

Schmidt, P. and C.A.K. Lovell. (1979). "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers." *Journal of Econometrics* 9, pp. 343–366.

Schmidt, P. and C.A.K. Lovell. (1980). "Estimating Stochastic Production and Cost Frontiers When Technical and Allocative Inefficiency are Correlated." *Journal of Econometrics* 13, pp. 83–100.

Stevenson, R.E. (1980). "Likelihood Functions for Generalized Stochastic Frontier Estimation." *Journal of Econometrics* 13, pp. 56–66.

Timmer, C.P. (1971). "Using a Probabilistic Frontier Function to Measure Technical Efficiency." *Journal of Political Economy* 79, pp. 776–794.