Sankhyā: The Indian Journal of Statistics
1999, Volume 61, Series B, Pt. 2, pp. 337-350

A MODEL OF FPF WITH CORRELATED ERROR COMPONENETS: AN APPLICATION TO INDIAN AGRICULTURE

By MANORANJAN PAL
and
ATANU SENGUPTA
Indian Statistical Institute, Calcutta

SUMMARY. The **S**tochastic **F**rontier **P**roduction **F**unction (**SFPF**) models usually assume that the two components of the equation error, namely the noise variable and the inefficiency error are independently distributed. In this paper we relax this assumption. This may serve as a basis for specification test. Moreover, the marginal distribution of the inefficiency error is taken as truncated normal instead of half normal. We develop a model of **SFPF** introducing both technical and allocative inefficiencies under the assumption that technical efficiency error and the noise variable are bivariate normal, the technical inefficiency being truncated at zero. The allocative inefficiency errors are also correlated among themselves. We obtain the likelihood function and discuss estimation of technical and allocative inefficiencies. This is illustrated through an Indian data on agricultural farms.

1. **Introduction**

In the standard neoclassical theory of firm behaviour, firms are generally assumed to optimize certain objective functions like profit, cost, revenue, sales or other such relevant economic variables. In this respect, a firm is adjudged to be efficient if the subsequent optimality conditions are satisfied. In reality, however, this is rarely the case. All such departures from the standard optimality conditions may be said to constitute inefficiency. The idea to measure inefficiency of firms in such cases has been introduced by Farrell (1957). Farrell defined inefficiency with respect to two attributes: Technical and Allocative. A firm is said to be technically inefficient (output inefficient) if it is possible to increase total output without altering the current levels of input-use. Similarly, an

Paper received. January 1998; revised March 1999.
AMS (1991) *subject classification.* 90A11, 90A19.
Key words and phrases. Stochastic frontier production function, technical and allocative efficiencies, correlated error components.

allocatively inefficient firm incurs a higher total input-cost because of wrong choice of proportions of inputs for a given level of output, the firm being a price-taker.

Ever since the introduction of the Frontier approach by Farrell for measuring productive efficiencies, a number of studies have been made in this direction (Fϕrsund, Lovell and Schmidt (1980), Bauer (1990)). There are mainly two competing lines of research to find out efficiencies. The first one is a non-parametric deterministic approach using mathematical programming which was later termed as **D**ata **E**nvelopment **A**pproach (**DEA**). The second one is a parametric statistical approach, which may be termed as **E**conometric **A**pproach (**EA**). There are advantages and disadvantages in either approach. The estimates obtained from the mathematical programming approach are highly sensitive to **E**rrors-in-**V**ariables (**EIV**s) especially for extreme observations. On the other hand, if the underlying distributional assumptions of the error terms in the equations are more or less valid and the mathematical forms of the equations can reasonably accurately approximate the actual form then **EA** is supposed to give better result even if the **EIV**s are present. This is more so for stochastic frontier functions. The assumptions can be substantially relaxed for panel data.

The essential idea behind the Frontier Production Function (FPF) approach (e.g., Aigner *et al.* (1977), Meeusen and van den Broeck (1977)) is that the output of each firm is bounded above by a frontier which is defined as the locus of maximum possible output levels for different levels of input combinations. The frontier is said to be stochastic if we introduce noise variable besides the inefficiency errors. The error term (the equation error) in the econometric models of stochastic frontier is thus assumed to be composed of two additive terms: one term is the noise variable, the other part is due to inefficiency and takes only negative value. This gives rise to what is known as technical inefficiency. The allocative inefficiency cannot be introduced in a single equation model. For this, one has to take a system approach. The effort in this direction started with cost system (cost and factor demand equations) and analytical solutions were obtained for some (Kumbhakar (1989), Schmidt and Lovell (1979)). Other system approaches such as using profit and distance functions were also tried. (Ali and Flinn (1983), Kumbhakar (1987), Ferrier and Lovell (1987)).

It has all along been argued in the literature of efficiency estimation that these two components of the error terms are independent. But this need not be true. For example, managerial efficiency may be affected by natural factors such as climatic conditions which is a statistical noise. It is easy to bring such an assumption in the context of time-series by arguing that current managerial decisions are affected by past natural shocks. Nonetheless if the lag of this effect is small, even in cross-sectional studies the effect may become important. In the context of Indian agriculture, where dependence on climatic factors is very high, the argument is reinforced further. In our exercise, for example, we have considered the cultivation of the paddy crop. In the region under study paddy is generally cultivated more than once a year (normally referred to as Aman, Boro and Aus). Naturally then a disturbance or any other incidence in any one of these seasons will affect decision making regarding cultivation in the subsequent seasons. Thus when we consider the cultivation of paddy as a whole (i.e., for the whole year) the inefficiency term and the statistical error term may and are likely to be correlated.

Another reason for such correlation may be due to misspecification error. There can exist some factors which influence the inefficiency error but are not included in the model (eg., personal bias, family norms and standards etc.). These factors are not often well defined or there may not be adequate data for them to be usable. Since the statistical error term is an impromptu variable, it contains all that is not included in the model. Hence it is necessary to test whether the two components of the error are correlated.

In this paper we develop a model of stochastic **FPF** introducing both technical and allocative inefficiency under the assumption that technical efficiency error is **correlated** with the noise variable.

The paper is divided into four sections. In section 2 we introduce our basic econometric model and develop the likelihood function. In section 3 we discuss how the technical and allocative inefficiencies are computed. Section 4 illustrates it through a data on Indian agricultural farms and makes some concluding observations.

2. The Basic Econometric Model

In this exercise we use a simple Cobb-Douglas production technology of the form

$$y \ = \ A(\prod_i x_i^{\alpha_i}) \exp(v) \qquad \ldots(1)$$

where y is the level of output produced, x_is are inputs $i(i = 1, 2, \ldots, n), v$ is the general statistical noise capturing random exogenous shocks and A is the technical efficiency parameter of the firm, α_is are the production parameters common to all firms.

In order to specify the stochastic nature of A, Kumbhakar (1987) suggested the following functional form as

$$A = \beta_0 \exp(\tau), \ \ \tau \leq 0 \qquad \ldots(1')$$

where τ represents technical efficiency which differs from firm to firm, while β_0 is common to all firms. Substituting $(1')$ in (1) we get

$$y \ = \ \beta_0 \prod x_i^{\alpha_i} \exp(\tau + v) \qquad \ldots(2)$$

The firm is said to maximize the profit function defined as

$$\Psi = Py - w'x \qquad \ldots(3)$$

where Ψ is the level of profit, P is the output price and w is the vector of input prices. The allocative inefficiency is then captured by the first order conditions of maximization as

$$\frac{w_i}{P} = mp_i exp(u_i) \qquad \ldots(4)$$

where $\exp(u_i)$ has been introduced to capture the effect of allocative inefficiency. mp_i is the marginal product of the input i .

In standard literature it is assumed that τ and v are independently distributed. This assumption is relaxed here. We also generalize the model by taking truncated distribution of the technical inefficiency component τ instead of taking half-normal. Symbolically,

$$\begin{array}{ll}(a) & \mathbf{u} \sim iid\ N_n(0, \Sigma), \\ (b) & v, \tau \sim\ Truncated\ N_2(\mu_\tau, \mu_v, \sigma_\tau, \sigma_v, \rho), \\ & -\infty < u_i < \infty \quad -\infty < v < \infty \quad \& \quad -\infty < \tau < h \quad . \end{array}$$

The density function of (v, τ) is given by

$$\begin{aligned} f(v,\tau) &= \frac{\Phi^{-1}(\frac{h-\mu_\tau}{\sigma_\tau})}{2\pi\sigma_v\sigma_\tau\sqrt{1-\rho^2}} \exp[\frac{-1}{2(1-\rho^2)}\{(\frac{\tau-\mu_\tau}{\sigma_\tau})^2 + (\frac{v-\mu_v}{\sigma_v})^2 \\ &\quad -2\rho(\frac{\tau-\mu_\tau}{\sigma_\tau})(\frac{v-\mu_v}{\sigma_v})\}], \quad -\infty < \tau < h, \quad -\infty < v < \infty. \end{aligned} \qquad \ldots(5)$$

Except for the term $\Phi^{-1}(\frac{h-\mu_\tau}{\sigma_\tau})$ which is coming due to truncation it is same as the bivariate normal density function.

Since $\mathbf{u}$ is independent of (v, τ), the joint density function is simply the product of the two density functions

$$\begin{aligned} f(v,\tau,\mathbf{u}) &= \frac{\Phi^{-1}(\frac{h-\mu_\tau}{\sigma_\tau})}{(2\pi)^{\frac{n+2}{2}}\sigma_v\sigma_\tau|\Sigma|^{1/2}\sqrt{1-\rho^2}} \exp[\frac{-1}{2}\{\frac{1}{(1-\rho^2)}\{(\frac{\tau-\mu_\tau}{\sigma_\tau})^2 + (\frac{v-\mu_v}{\sigma_v})^2 \\ &\quad -2\rho(\frac{\tau-\mu_\tau}{\sigma_\tau})(\frac{v-\mu_v}{\sigma_v})\} + \mathbf{u}'\Sigma^{-1}\mathbf{u}\}] \\ &\quad -\infty < \tau < h, \quad -\infty < v < \infty, \quad \infty < u_i < \infty, \quad \forall i = 1, \ldots, n. \end{aligned} \qquad \ldots(6)$$

The relations between $v, \tau \& \mathbf{u}$ with the observations $y, x_1, x_2, \ldots, x_n$ are as follows (Kumbhakar (1987))

$$\ln y_f = \ln \beta_0 + \sum_{i=1}^{n} \alpha_i \ln x_{if} + \tau_f + v_f$$

and

$$\ln(\frac{w_{if}}{P_f}) - \ln \beta_0 - \ln \alpha_i + \ln x_i - \sum_{k=1}^{n} \alpha_k \ln x_{kf} = \tau_f + u_{if} \quad f = 1, 2, \ldots, F \qquad \ldots (7)$$

where F is the number of firms.

The data for our model are $y_f, x_{1f}, \ldots, x_{nf}, w_{1f}, \ldots, w_{nf}$ and P_f for $f = 1, 2, \ldots, F$. and the suffix f denotes the specific firm. We shall henceforth suppress the firm specific subscripts.

Substituting $z_1 = v + \tau, \mathbf{z_2} = \mathbf{u} + l\tau$ and $z_3 = \tau$ where $l = (1, 1, \ldots, 1)'$ and noting that Jacobian of transformation is one, we get

$$f(z_1, \mathbf{z_2}, z_3) = \frac{\Phi^{-1}(\frac{h - \mu_\tau}{\sigma_\tau})}{(2\pi)^{\frac{n+2}{2}} \sigma_v \sigma_\tau |\Sigma|^{1/2} \sqrt{1 - \rho^2}} \exp[\frac{-1}{2}\{(\frac{z_3 - \mu_*}{\sigma_*})^2 + a_*\}]$$

where

$$\begin{aligned}
\frac{1}{\sigma_*^2} &= \frac{1}{(1-\rho^2)}(\frac{1}{\sigma_\tau^2} + \frac{1}{\sigma_v^2} + \frac{2\rho}{\sigma_\tau \sigma_v}) + (l'\Sigma^{-1} l) \\
\mu_* &= \sigma_*^2[\frac{1}{1-\rho^2}\{\frac{(z_1 - \mu_v)}{\sigma_v^2}(1 + \rho\frac{\sigma_v}{\sigma_\tau}) + \frac{\mu_\tau}{\sigma_\tau^2}(1 + \rho\frac{\sigma_\tau}{\sigma_v})\} + (l'\Sigma^{-1}\mathbf{z_2})] \\
a_* &= [\frac{1}{1-\rho^2}\{\frac{\mu_\tau^2}{\sigma_\tau^2} + (\frac{z_1 - \mu_v}{\sigma_v})^2 + 2\rho(\frac{\mu_\tau}{\sigma_\tau})(\frac{z_1 - \mu_v}{\sigma_v})\} + (\mathbf{z_2'}\Sigma^{-1}\mathbf{z_2})] \\
&\quad -\sigma_*^2[\frac{1}{1-\rho^2}\{\frac{(z_1 - \mu_v)}{\sigma_v^2}(1 + \rho\frac{\sigma_v}{\sigma_\tau}) + \frac{\mu_\tau}{\sigma_\tau^2}(1 + \rho\frac{\sigma_\tau}{\sigma_v})\} + l'\Sigma^{-1}\mathbf{z_2}]^2 \\
&\quad -\infty < z_1 < \infty, \quad (-\infty)^n < \mathbf{z_2} < (\infty)^n \quad \& -\infty < z_3 < h.
\end{aligned}$$

The density function of $(z_1, \mathbf{z_2})$ can be easily obtained by integrating $f(z_1, \mathbf{z_2}, z_3)$ over z_3 in the appropriate range

$$\begin{aligned}
f(z_1, \mathbf{z_2}) &= \int_{-\infty}^{h} f(z_1, \mathbf{z_2}, z_3) dz_3 \\
&= \frac{\sigma_* exp(-\frac{a_*}{2})}{(2\pi)^{\frac{n+1}{2}} \sigma_v \sigma_\tau |\Sigma|^{1/2} \sqrt{1-\rho^2}\Phi(\frac{h-\mu_\tau}{\sigma_\tau})} \Phi(\frac{h - \mu_*}{\sigma_*}) \qquad \ldots (8)
\end{aligned}$$

To get the density function of the observed values namely $y_f, x_{1f}, \ldots, x_{nf}$ one has to multiply this by the relevant Jacobian of transformation $(1 - \sum \alpha_i)$. The likelihood function is obtained by multiplying these density functions for all observations and writing it as a function of parameters $(\beta_0, \alpha, \sigma_v^2, \sigma_\tau^2, \rho, \Sigma, \mu_\tau)$. The log likelihood function, by introducing firm specific suffix, is

$$\begin{aligned} L(\beta_0, \alpha, \sigma_v^2, \sigma_\tau^2, \rho, \Sigma, \mu_\tau) &= -\frac{F(n+1)}{2}\ln(2\pi) - \frac{F}{2}\ln\sigma_v^2 - \frac{F}{2}\ln\sigma_\tau^2 \\ &\quad -\frac{F}{2}\ln|\Sigma| + \frac{F}{2}\ln\sigma_*^2 - \frac{1}{2}\sum_f a_{*f} \\ &\quad -F\ln\Phi(\frac{h-\mu_\tau}{\sigma_\tau}) + \sum_f \ln\Phi(\frac{h-\mu_{*f}}{\sigma_*}) \\ &\quad -\frac{F}{2}\ln(1-\rho^2) \end{aligned}$$

3. Estimation of Efficiencies

The log-likelihood function can be maximized to estimate the parameters of the model. (But there may be identification problem if μ_v is not taken as zero. Also since the technical efficiencies are supposed to be less than or equal to one we must assume $h = 0$). The next task is to estimate inefficiency errors namely the technical inefficiency error τ_f and the allocative inefficiency errors u_{if}'s for each firm f. From the model however we can only estimate $\tau_f + v_f$ and $\tau_f + u_{if}$. There is no way of separating these components. Jondrow *et al.* (1982) suggested use of conditional expectation of τ_f given $\tau_f + v_f$ to get an indirect estimate of τ_f for the technical efficiency. We can similarly find the conditional expectation of u_{if} given $(\tau_f + v_f, \tau_f + u_{if}, \ldots, \tau_f + u_{nf})$. One need not confine oneself to take expectation only. Any other measure of central tendency like mode will serve the purpose. In our analysis we shall take conditional modes.

We take the conditional distribution of z_3 given z_1 and $\mathbf{z_2}$ to get the distribution of the inefficiency term $z_3(=\tau)$ for each observation following suggestion made by Jondrow et al.(1982).

$$\begin{aligned} f(z_3|z_1, \mathbf{z_2}) &= \frac{f(z_1, \mathbf{z_2}, z_3)}{f(z_1, \mathbf{z_2})} \\ &= \frac{\exp[\frac{-1}{2}(\frac{z_3-\mu_*}{\sigma_*})^2]}{\sqrt{2\pi}\sigma_*\Phi(\frac{h-\mu_*}{\sigma_*})} \end{aligned}$$

It is nothing but the density function of a normal distribution with mean μ_* and variance σ_*^2 truncated at h from above. The mode and expectation of the conditional distribution which may be taken to find a measure of technical efficiency (the subscription f is compressed) is given as

$$M_0(z_3|z_1,\mathbf{z_2}) = \begin{cases} \mu_* \; if \; \mu_* \leq h \\ h \; if \; \mu_* > h \end{cases}$$

$$E(z_3|z_1,\mathbf{z_2}) = \mu_* - \sigma_* \frac{\phi(\frac{h-\mu_*}{\sigma_*})}{\Phi(\frac{h-\mu_*}{\sigma_*})}$$

The allocative inefficiency is found by taking conditional distribution of u_j given z_1 and $\mathbf{z_2}$ for all j and for each observation. For this we take the following transformation of v, τ and $\mathbf{u}$.

$$z_1 = v + \tau, \quad \mathbf{z_2} = \mathbf{u} + l\tau \quad \text{and} \quad Z_3 = u_j.$$

Thus,

$$u_j = Z_3, \quad \tau = z_{2j} - Z_3 \quad \text{and} \quad \mathbf{u} = \mathbf{z_2} - l(z_{2j} - Z_3),$$

$$-\infty < z_1 < \infty, \qquad (-\infty)^n < \mathbf{z_2} < (\infty)^n \quad \text{and} \quad z_{2j} - h < Z_3 < \infty.$$

The Jacobian of transformation being 1, we have

$$f(z_1,\mathbf{z_2},Z_3) = \frac{\Phi^{-1}(\frac{h-\mu_\tau}{\sigma_\tau})}{(2\pi)^{\frac{n+2}{2}}\sigma_v\sigma_\tau|\Sigma|^{1/2}\sqrt{1-\rho^2}} \exp[-\frac{1}{2}(\frac{Z_3-\mu_{**}}{\sigma_{**}})^2]\exp(-\frac{a_{**}}{2}),$$

where

$$\begin{aligned}
\frac{1}{\sigma_{**}^2} &= \frac{1}{(1-\rho^2)}(\frac{1}{\sigma_\tau^2} + \frac{1}{\sigma_v^2} + \frac{2\rho}{\sigma_\tau\sigma_v}) + (l'\Sigma^{-1}l) \\
\mu_{**} &= \sigma_{**}^2\{\frac{1}{1-\rho^2}[\frac{z_{2j}-\mu_\tau}{\sigma_\tau^2}(1+\rho\frac{\sigma_\tau}{\sigma_v}) - \frac{z_1-z_{2j}-\mu_v}{\sigma_v^2}(1+\rho\frac{\sigma_v}{\sigma_\tau})] - l'\Sigma^{-1}(\mathbf{z_2}-lz_{2j})\} \\
a_{**} &= \frac{1}{1-\rho^2}[(\frac{z_{2j}-\mu_\tau}{\sigma_\tau})^2 + (\frac{z_1-z_{2j}-\mu_v}{\sigma_v})^2 - \frac{2\rho}{\sigma_v\sigma_\tau}(z_{2j}-\mu_\tau)(z_1-z_{2j}-\mu_v)] \\
&\quad +(\mathbf{z_2}-lz_{2j})'\Sigma^{-1}(\mathbf{z_2}-lz_{2j}) - \\
&\quad \sigma_{**}^2\{\frac{1}{1-\rho^2}[\frac{z_{2j}-\mu_\tau}{\sigma_\tau^2}(1+\rho\frac{\sigma_\tau}{\sigma_v}) - \frac{z_1-z_{2j}-\mu_v}{\sigma_v^2}(1+\rho\frac{\sigma_v}{\sigma_\tau})] - l'\Sigma^{-1}(\mathbf{z_2}-lz_{2j})\}^2,
\end{aligned}$$

so that we get

$$\begin{aligned}
f(z_1,\mathbf{z_2}) &= \int_{z_{2j}-h}^{\infty} f(z_1,\mathbf{z_2},Z_3)dZ_3 \\
&= \frac{\sigma_{**}exp(-\frac{a_{**}}{2})}{(2\pi)^{\frac{n+1}{2}}\sigma_v\sigma_\tau|\Sigma|^{1/2}\sqrt{1-\rho^2}\Phi(\frac{h-\mu_\tau}{\sigma_\tau})}(1-\Phi(\frac{z_{2j}-h-\mu_{**}}{\sigma_{**}})).
\end{aligned} \quad \ldots(9)$$

One can see that the expression of $f(z_1, \mathbf{z_2})$ in (8) is not same as that written in (9). It can however be shown that they are equivalent noting that

$$\sigma_{**} = \sigma_*$$

$$z_{2j} - \mu_{**} = \mu_*$$

$$a_{**} = a_*$$

and

$$1 - \Phi(\frac{z_{2j} - h - \mu_{**}}{\sigma_{**}}) = \Phi(\frac{h - (z_{2j} - \mu_{**})}{\sigma_{**}}).$$

The conditional distribution of Z_3 given z_1 and $\mathbf{z_2}$ is

$$\begin{aligned} f(Z_3|z_1, \mathbf{z_2}) &= \frac{f(z_1, \mathbf{z_2}, Z_3)}{f(z_1, \mathbf{z_2})} \\ &= \frac{\exp[\frac{-1}{2}(\frac{Z_3 - \mu_{**}}{\sigma_{**}})^2]}{\sqrt{2\pi}\sigma_{**}(1 - \Phi(\frac{z_{2j} - h - \mu_{**}}{\sigma_{**}}))}. \end{aligned}$$

It is again a normal distribution with mean μ_{**} and variance σ^2_{**} truncated at $z_{2j} - h$ from below. The mode and expectation of this distribution (the subscription f is compressed) is similarly derived as

$$M_0(Z_3|z_1, \mathbf{z_2}) = \begin{cases} \mu_{**} & if \ z_{2j} - h \leq \mu_{**} \\ z_{2j} - h & if \ z_{2j} - h > \mu_{**} \end{cases}$$

$$E(Z_3|z_1, \mathbf{z_2}) = \mu_{**} + \sigma_{**} \frac{\phi(\frac{z_{2j} - h - \mu_{**}}{\sigma_{**}})}{(1 - \Phi(\frac{z_{2j} - h - \mu_{**}}{\sigma_{**}}))} \quad .$$

4. An Illustration

4.1 *The data.* The data which we have used in this exercise were collected by the Ministry of Agriculture, Government of India through the "Comprehensive Scheme for Studying Cost of Cultivation" (CSSCC). The data were collected for every year beginning in 1971 from various parts of India. We have used in this study farm-level disaggregate data pertaining to the year 1989-90 for West Bengal.

For the purpose of collecting CSSCC data in West Bengal, the entire state was divided into six agro-climatic zones based on cultivation practices, type of soil, irrigation facilities and rainfall. Considering the information regarding total cropped area under Aus Paddy, Aman Paddy, Boro Paddy, Jute, Potato and Wheat, zonal allocation of sixty blocks was made.

A multistage random sampling design was adopted from blocks to mouza and then from mouza to households. The landless labourers were excluded from the set of households. In this way a total of 600 households were selected. Out of these 600 households, only 597 cultivates paddy. Our study is based on these 597 paddy cultivating households only. After scrutiny we were left with only 457 households as we had to delete households with outlying observations and/or with some variables missing.

This data set supplies information on various inputs like human labour, bullock labour, fertilizer, manure, machine and output of paddy both in value and quantitative terms. For our illustration we have taken only three inputs namely (i) human labour hour, (ii) bullock hour and (iii) fertilizer, which presumably explain production of paddy very well. All these variables are measured in per unit area. The time period is one year. Since quantities and values are both available we can obtain price data. The prices will be different for different farms because of variations in quality. A total of 457 farms were taken for our study. In table 1 we provide the summary statistics for the inputs and output.

Table 1. Summary description of the variables used

	Mean Values			
	Labour hour	Bullock Hour	Fertlizer	Output of Paddy
Quantity	1248.6	214.8	74.7	32.24
Price	2.10	3.99	5.43	266.43
	Variances			
Quantity	131262.3	14675.1	3547.9	1154.4
Price	0.227	4.856	0.499	1154.405

For convenience we shall write the quantities as x_1, x_2, x_3 and the corresponding prices as w_1, w_2, w_3. The input is denoted as y and its price as P. In most of the cases the correlation between quantities and prices have been found to be negligible. The correlation matrix is as follows:

$$\begin{pmatrix} & \text{P} & w_1 & w_2 & w_3 \\ \text{y} & \text{-0.011} & \dots & \dots & \dots \\ x_1 & \dots & 0.094 & \dots & \dots \\ x_2 & \dots & \dots & \text{-0.032} & \dots \\ x_3 & \dots & \dots & \dots & \text{-0.247} \end{pmatrix}$$

4.2 *Initial estimates.* We have applied the S-Plus package for maximizing the likelihood function. In fact we have minimised the negative of the log likelihood function. For this it was necessary to supply the initial values of the parameter vector namely

$$\vec{p} \equiv (\beta(0), \alpha_1(0), \alpha_2(0), \alpha_3(0), \sigma_v(0), \sigma_\tau(0), \rho(0), \sigma_{11}(0), \sigma_{12}(0), \sigma_{13}(0), \sigma_{22}(0), \sigma_{23}(0), \sigma_{33}(0), \mu_\tau(0))$$

There are at least two reasons for selecting the vector $\vec{p}$. Firstly this would guarantee convergence. Moreover it would save a lot of computing time. Since we took in our study an arbitrary initial values. It took about 300 iterations involving 5 hours of computing time in a Pentium using S-plus package. We may now discuss various ways of ascertaining initial estimates.

The factor cost share equations (7) may be written as

$$ln\ (w_i x_i) = ln\ \alpha_i + u_i + ln(Py) \qquad \ldots (10)$$

Since u_i is normal with mean zero, a good estimate of $ln\ \alpha_i$ would be

$$\widehat{ln\ \alpha_i} = \overline{ln\ (w_i x_i)} - \overline{ln\ (Py)}.$$

Again from (10), we have $\dfrac{w_i x_i}{Py} = \alpha_i e^{u_i}$. or

$$\begin{aligned} \left(\overline{\frac{w_i x_i}{Py}}\right) &\simeq \alpha_i\ E(e^{u_i}) = \alpha_i E(1 + u_i + \frac{u_i^2}{2} + \ldots) \\ &\simeq \alpha_i (1 + \frac{\sigma_{ii}}{2}) \end{aligned}$$

Thus, an initial estimate of σ_{ii} would be

$$\sigma_{ii}(0) = 2 \left\{ \frac{1}{\alpha_i(0)} \left(\overline{\frac{w_i x_i}{Py}}\right) - 1 \right\}.$$

Finding initial estimates of σ_v^2 and σ_τ^2 are difficult. We may assume $\sigma_v^2 = \sigma_\tau^2$ and also $\rho(v, \tau) = 0$. Since

$$y = \beta_0 \Pi\ x_i^{\alpha_i}\ exp(\tau + v),$$

one may regress $ln\ y$ on $ln\ x_i$'s to get the residual variance $\sigma_e^2(0)$, say. Hence

$$\sigma_v^2(0) = \sigma_\tau^2(0) = \sigma_e^2(0)/2.$$

Also, the intercept from the same regression can be used to get an initial estimate of $ln\beta_0$. Since $\tau \leq 0$, initial estimate of $ln\ \beta_0$ should be greater than the value of the intercept. All correlations were initially taken as zeroes by us.

4.3 *The results and the concluding remarks.* The frequency distributions of z_1 (estimating $(v + \tau)$) and z_2 (estimating $(u + l\tau)$) together with their histograms clearly indicate that the distributions are negatively skewed. This is expected as τ was included in both the error terms. Distribution of μ_* (Table 2) and μ_{**} (Table 4) are similarly shaped. Though the correlation between v and τ seems to be low, we have found the value of ρ to be highly significant. This only proves that it is sometimes necessary to have the assumption of correlation between technical error and noise variable in the model. The covariances between u_is are also rather small. However we have not tested whether these are significantly different from zero.

Table 2. FREQUENCY DISTRIBUTION OF μ_*

Levels of μ_*	Frequency	Percentage of total
Below -0.5	6	1.3
-0.5 to -0.4	24	5.3
-0.4 to -0.3	51	11.2
-0.3 to -0.2	72	15.8
-0.2 to -0.1	90	19.7
-0.1 to 0	105	23.0
0	109	23.7
ALL	457	100

Table 3. FREQUENCY DISTRIBUTION OF THE LEVELS OF TECHNICAL EFFICIENCY

Efficiency Levels	Frequency	Percentage of total
0-0.5	1	0.2
0.5-0.6	5	1.1
0.6-0.7	47	10.3
0.7-0.8	69	15.1
0.8-0.9	119	26.0
0.9-1.0	107	23.4
1.0	109	23.9
ALL	457	100

Table 4. FREQUENCY DISTRIBUTION OF μ_{**}

Levels of μ_{**}s	Frequency Distribution of		
	μ_{**l}	μ_{**b}	μ_{**f}
Below -1.5	1	27	38
	(0.2)	(6.0)	(8.3)
-1.5 to -1	6	29	59
	(1.3)	(6.3)	(12.7)
-1 to -0.5	69	77	57
	(15.1)	(16.8)	(12.5)
-0.5 to 0	248	94	93
	(54.3)	(20.6)	(20.4)
0 to 0.5	123	118	102
	(26.9)	(25.8)	(22.3)
0.5 to 1	10	86	88
	(2.2)	(18.8)	(19.3)
1 & Above	0	26	20
	(0)	(5.7)	(4.5)
ALL	457	457	457
	(100)	(100)	(100)

μ_{**l} represents the value of μ_{**} for human labour
μ_{**b} represents the value of μ_{**} for bullock labour
μ_{**f} represents the value of μ_{**} for fertilizer

The technical and allocative efficiency parameters are calculated using the methodology described in the text. The mean technical efficiency (Table 3) is rather high 0.871 with standard deviation 0.121. It is found that there is very little dependence between size and technical efficiency. Size does not seem to influence the pattern of technical efficiency.

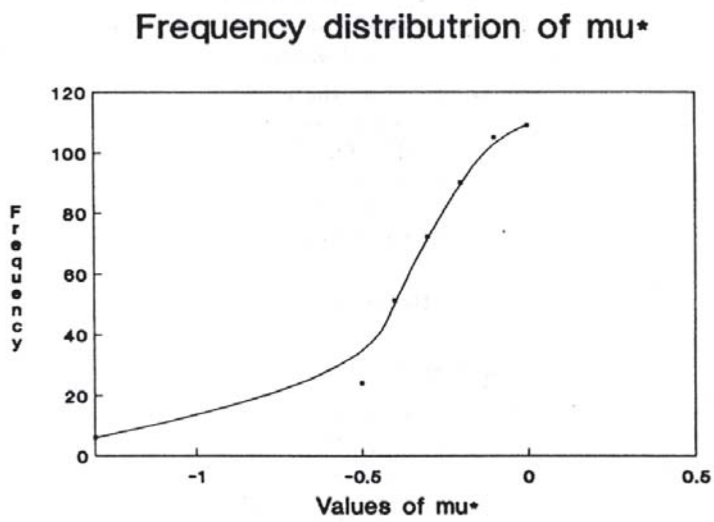
As for allocative inefficiency of labour hour (Table 5) the mean allocative inefficiency is 0.990 with standard deviation 0.251. For bullock labour mean inefficiency turns out to be 1.371 and standard deviation is 1.108. Fertilizer has a mean inefficiency of 1.184 and standard deviation of 0.791. On average labour hour has turned out to be most efficient with least relative variation in efficiency (coefficient of variation). From the coefficients of the inputs it is clear that labour accounts for about 29% of the cost incurred while the total input cost share is about 40%.

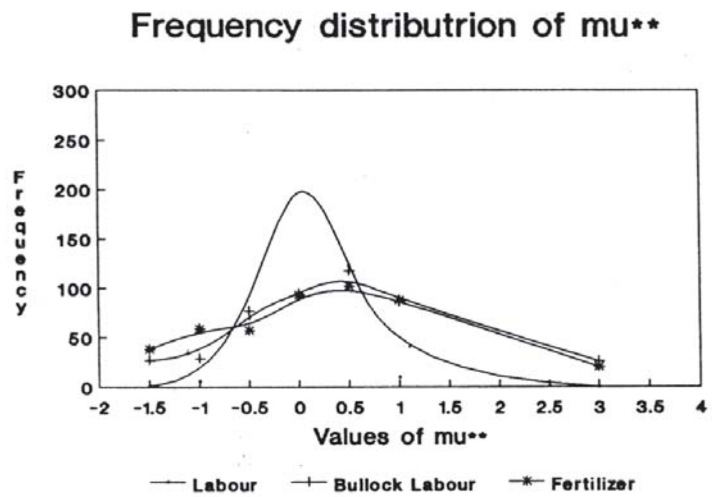
Table 5. Frequency distribution of the levels of Allocative Inefficiency

Levels	Frequency Distribution of		
	ae_l	ae_b	ae_f
0-0.5	10	75	108
	(2.2)	(16.4)	(23.6)
0.5-0.6	15	28	18
	(3.3)	(6.1)	(3.9)
0.6-0.7	24	21	22
	(5.3)	(4.6)	(4.8)
0.7-0.8	47	24	30
	(10.3)	(5.3)	(6.6)
0.8-0.9	94	36	27
	(20.6)	(7.9)	(5.9)
0.9-1.0	79	20	20
	(17.3)	(4.4)	(4.4)
1-1.5	176	101	98
	(38.5)	(22.1)	(21.4)
1.5-2	11	63	65
	(2.4)	(13.8)	(14.2)
2 & Above	1	89	69
	(0.2)	(19.4)	(15.2)
ALL	457	457	457
	(100)	(100)	(100)

ae_l, ae_b, ae_f represents the value of allocative inefficiency for human labour, bullock labour and fertilizer respectively

In this paper the generalizations of the **FPF** have been made in two directions. First, the inefficiency error has been assumed to be correlated with the noise variable. Second, the inefficiency error is said to follow a truncated normal instead of half-normal distribution. For this it is assumed that the joint distribution follows a bivariate normal, the inefficiency error being truncated at some known point zero (say). The normality assumption of inefficiency error may


Frequency distributrion of mu*
Frequency
120
100
80
60
40
20
0
-1
-0.5
0
0.5
Values of mu*


Frequency distributrion of mu**
Frequency
300
250
200
150
100
50
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
Values of mu**
Labour
Bullock Labour
Fertilizer

not be appropriate for some situations. We can take any other distribution truncated at zero from above or we may take the negative of gamma, lognormal, beta or some related distributions. To introduce non-independence, one may assume the noise variable v, conditional to the inefficiency error τ, to be normally distributed with mean $\rho\frac{\sigma_v}{\sigma_\tau}(\tau - \mu_\tau)$ and variance $\sigma_v^2(1-\rho^2)$ where μ_τ is the expectation and the ρ is interpreted as the correlation coefficient. The joint distribution can be simply computed by multiplying the conditional density function with the density function of τ. One can see that this is an extension of the model discussed in the paper, because, just by assuming τ to be normal it reduces to the earlier model. The log likelihood function of this model and the technical and allocative efficiencies can be calculated in the similar manner.

References

AIGNER D.J., C.A.K.LOVELL AND P. SCHIMDT (1977), Formulation and estimation of frontier production models, *Jour. Econometrics* **6**(1) (July), 21-37.

ALI, M. AMD J.C. FLINN (1989), Profit efficiency among Basmati rice producers in Pakistan Punjab, *Amer. Jour. Agricultural Economics* **71**(2) (May), 303-10.

BAUER, P.W. (1990), Decomposing TFP growth in the presence of cost efficiency, nonconstant returns to scale, and technological progress, *Jour. Productivity Analysis* **1**(4) (June), 287-300.

FARELL M.J., (1957) The measurement of productive efficiency, *Jour. Royal Statist. Soc.* Series A, General, **120**(3), 253-81.

FERRIER, G.D. AND C.A.K. LOVELL (1990), Measuring cost efficiency in banking: econometric and linear programming evidence, in A.Y. Lewin and C.A.K. Lovell, eds. (1990), *Frontier Analysis: Parametric and Non-parametric Approaches, Jour. Econometrics* **46**(1/2) (October/November).

FϕRSUND, F.R., C.A.K. LOVELL AND P. SCHMIDT (1980), A survey of frontier production functions and of their relationship to efficiency measurement, *Jour. Econometrics* **13**, 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 model, *Jour. Econometrics* **19**(2/3) (August), 233-238.

KUMBHAKAR, S. (1987), The specification of technical and allocative inefficiency in stochastic production and profit frontiers *Journal of Econometrics* **34**(3), 335-348.

KUMBHAKAR, S.C. (1989), Modeling technical and allocative inefficiency in translog production function, *Economic Letters* **31**(2), 119-124

MEEUSEN, W. AND J. VAN DEN BROECK (1977), Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review* **18**(2), 435-444.

SCHIMDT, P. AND C.A.K. LOVELL (1979), Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers, *Jour. Econometrics* **9**(3), 343-366.

MANORANJAN PAL AND ATANU SENGUPTA
ECONOMIC RESEARCH UNIT
INDIAN STATISTICAL INSTITUTE
203 B.T. ROAD
CALCUTTA-700035
INDIA
e-mails: mano@isical.ac.in/res9430@isical.ac.in