Recent Developments in Stochastic Frontier Model with Correlated Error Components

Arabinda Das

Abstract:

The possibility of measuring efficiency of a firm on the basis of information on inputs and outputs was first introduced by Farrell (1957), followed by Koopmans (1951) and Debreu (1951). Since then one of the most widely used approach of efficiency analysis has been frontier approach. Aigner and Chu (1968) suggested a deterministic frontier model for estimating inefficiency which later found to have limitation of isolation of inefficiency from the noise as both are mixed together in the error term of the model. This model was improved with inclusion of noise, termed as stochastic frontier model (SFM), by Aigner, Lovell and Schmidt (1977), Meeusen and van den Broeck (1977) almost simultaneously where the error is decomposed into two random components representing the noise and inefficiency with the assumption of independence. After a long period researchers argued the validation of the independence assumption between the error components. In recent years several researches have been developing SFM with correlated error components under different modelling approaches. This paper surveys this journey of SFM with correlated error components and discusses different statistical approaches already developed to model SFM with correlated error components.

Key words: Frontier approach, deterministic frontier model, stochastic frontier model, copula function.

JEL Classification: C3, C10, C52

1. Introduction

Ever since the stochastic frontier model (SFM) was proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977), the last three decades has seen extensive development in the literature of the SFM. The central research of the SFM is to estimate the inefficiency of a concerned firm or economic agent. Two broad paradigms for measuring inefficiency developed throughout this period, one based on an econometric approach to estimation of theory based models of production, cost or profit and other based on nonparametric, programming approach to analysis of observed outcomes. In this paper we survey the underlying models and econometric techniques that have been used in studying technical inefficiency in the stochastic frontier framework, on which a number of review papers (see: among others, Greene, 1993 and Murillo-Zamorano, 2004), and a few books (Coelli, Rao and Battese, 1998; Fried, Lovell and Schmidt, 1993 and Kumbhakar and Lovell, 2000) have already been written. We also present some of the recent developments in econometric methodology in the context of SFM with correlated error components. Needless to say the present review is not comprehensive and owes much to these excellent books and review articles. The omission of other topics does not mean that we consider them unimportant. Also, the author restricts the survey within the cross-section data model only.

Many applications of the SFM in diverse field have been coming up since its inception. SFM has many applications in management discipline like financial industries and resource-based industries. We have reported a few of them here. Khatri et al. (2002) estimated inefficiency for corporate sector performance and examined the role of corporate governance for a panel dataset of 31 of the largest non-financial companies listed on the Kuala Lumpur stock exchange for the period 1995 to 1999. Dawson & Dobson (2002) used SFM to measure managerial efficiency for English football association. Troutt et al. (2005) analysed the performance of a set of mutual funds from morning star database using SFM. Camanho and Dyson (2005) estimated the cost efficiency for a set of bank branches. Cummins et al. (2005) measured the efficiency for the Spanish insurance industry. Silva et al. (2005) measured profit efficiency for a set of Portuguese bank branches and found that profit improvement of bank branches is due to allocative efficiency in the long run whereas technical efficiency is responsible for the short run. Dar A. A. (2007) studies firm efficiency in studies of labor market of Canada. Dolton et al. (2007) examined how students allocated their time to efficiently score from a survey conducted in April 1999 on first and final year students from the different qualifications offered at the University of Malaga. Choi, K.-S. (2011) used stochastic frontier function to measure the efficiency of baseball manager. Brissimis et al. (2011) estimated efficiency on a panel dataset of commercial banks of 13 EU countries for the period 1996 to 2003 and found that the most technically efficient banking sectors were in Austria, Germany and the UK.

The plan of the paper is as follows. Section 2 describes modeling of technical inefficiency which is the prime interest of these models, and Section 3 considers econometric analysis of estimating the technical inefficiency. Section 4 describes the basic cross-sectional SFM, its different methods of estimation. Section 5 discusses different approaches of modeling and estimation methods of SFM with correlated error components. In Section 6 we present an illustrative study of the different modeling approaches using US electricity data. Finally, Section 7 gives the concluding remarks and some future research developments in the field.

2 Modeling Inefficiency

The production frontier or its dual cost frontier, or the convex conjugate of the two, the profit frontier, represents the maximum output obtained with a given set of inputs or the minimum cost of producing that output given the prices of the inputs or the maximum profit attainable given the output, inputs, and prices of the inputs. The estimation of frontier functions is the econometric exercise with the underlying theoretical proposition that no observed firm can exceed the frontier. Measurement of (in)efficiency is, then, the empirical estimation of the extent to which observed agents (fail to) achieve the theoretical frontier. The estimated model of production, cost or profit is the means to the objective of measuring inefficiency.

The measures of the different concepts of efficiency, addressed by Debreu (1951), Farrell (1957), Koopmans (1951) and Shephard (1953) are based on the economic notion of 'technology'. A firm's technology can be alternatively represented by Production Set (T), Output Set $(P(x))$ or Input Set $(L(y))$.

Let a firm uses inputs $x = (x_1, ..., x_n) \in R_+^n$ to produce outputs $y = (y_1, ..., y_m) \in R_+^m$. A production technology is given by the input-output combination $z = (x, y)$. A production technology is said to be feasible if and only if y can be produced from the input x using the given technology i.e. $T = \{(y, x) = z : x \text{ can produce } y\} \subset R_+^{m+n}$.

The output set $P(x)=\{y:z=(y, x) \in T\} \subset R_+^m$ consists of all those feasible outputs that can be produced using the input x and the input set $L(y)=\{x:z=(y, x) \in T\} \subset R_+^n$ consists of all those feasible inputs x that can produce a given level of output y.

Shephard (1953, 1970) introduced the notion of distance function to provide a functional form of technology. Distance function can be described in terms of both the output set $P(x)$ and the input set $L(y)$. The input and output distance functions are respectively $d_i(y,x)= max(\lambda:x/\lambda \in L(y), \lambda>1)$ and $d_0(y,x)=max(\mu:y/\mu \in P(x), 0<\mu<1)$. This input distance function gives the maximum radial reduction in input to produce a given level of output whereas output distance function gives the maximum radial expansion of output possible for a given level of inputs.

The Debreu - Farrell input and output oriented measures of technical efficiency are respectively $TE_i(y, x)= max(\theta : \theta x \in L(y))$ and $TE_0(x, y)=max(\phi : \phi y \in P(x))$.

Let a firm uses inputs x with prices $q = (q_1....q_n)$ and we have single output y that can be sold at price p. The production frontier is then:

$$f(x)=max\{y:y \in P(x)\}=max\ \{y:x \in L(y)\}$$

which is the maximum possible output technologically feasible for this level of inputs.

In this case, the output oriented measure of technical efficiency becomes ratio of maximum to actual output

$$TE_0(x, y)=[max(\phi y \leq f(x)]^{-1}$$

3 Econometric Analysis of the Efficiency

Cobb and Douglas (1928) estimated production function using OLS technique using the observed outputs and inputs to estimate the efficiency of a firm in early twentieth century long before Farrell (1957). However, the assumption of OLS with observations lie around the estimated model violated the frontier property of the production function that each observation lies below the frontier. Next forty years ignored the frontier property of the production function and estimated 'average' production function which ruled out its use for efficiency estimation.

The possibility of measuring the efficiency of a firm on the basis of information on the inputs used and output produced started with Debreu (1951), Farrell (1957) and Koopmans (1951). Two different approaches of efficiency analysis have been developed since then: one is the Frontier Approach and the other is Data Envelopment Approach (DEA). Since the present work follows the frontier approach of efficiency, we omit the review of the vast literature on efficiency analysis by the Data Envelopment Approach. Detailed discussions on this approach of efficiency analysis can be found in Coelli, Rao and Battese (1998); Fried, Lovell and Schmidt (1993).

The frontier approach is based on the econometric analysis and requires stochastic specification of the frontier function. This approach estimates the frontier model and provides some estimators for efficiency. Estimation of efficiency under this approach is carried out on deterministic as well as stochastic frontier models. In a deterministic frontier model inefficiency is represented by a single one-sided stochastic term with no specification of noise whereas the stochastic frontier model has random noise separated out from inefficiency leaving the observational errors with two latent components called noise and inefficiency.

Consider a set of 'n' firms produces a single output using a certain technology and 'K' inputs. Let y_i be the output, and x_i be the vector of the inputs used, by the ith firm. Then the ith observational equation of the production frontier model is given by

$$y_i = f(x_i, \beta)TE_i, \ i = 1,2,....,n \qquad (3.1)$$

where $f(x_i, \beta)$ is the production frontier, β is the vector of unknown technological parameters and TE_i is the output oriented efficiency of ith firm. All the observations satisfy the frontier property with respect to the estimated production frontier. Hence, we have

$$TE_i = y_i/f(x_i, \beta) \qquad (3.2)$$

which defines the technical efficiency as the ratio of observed output to the frontier output under the current technology. The amount by which an observation lies below the frontier is called inefficiency when $TE_i < 1$. The production frontier model given in equation (3.1) is called deterministic frontier model. This model was estimated by Aigner and Chu (1968) using programming technique. Richmond (1974) improved upon the COLS estimates to make them unbiased and consistent. In order to give statistical content to the programming estimators proposed by Aigner and Chu (1968), Schmidt (1976) estimated the model (3.1) by the maximum likelihood (ML) method assuming exponential and half-normal distribution. Later, Greene (1980) estimated another deterministic frontier model assuming ε_i's are distributed as gamma variables. This model too was estimated by the ML method.

4 The Stochastic Frontier Model

Although the deterministic frontier approach of Aigner and Chu (1968) and Schmidt (1976) estimates the frontier function respecting its frontier property, an obvious limitation of this approach is that one cannot isolate the effect of inefficiency from that of the random noise as both are lumped together in the disturbance term of the model. Also, it violates one of the regularity conditions required for application of ML method viz. the support of the distribution of y must be independent of the parameter vector.

The stochastic frontier approach of efficiency analysis which aimed to rectify the above mentioned limitation of the deterministic frontier approach, was introduced by Aigner, Lovell and Schmidt (1977), Meeusen and van den Broeck (1977) almost simultaneously. The novelty of the stochastic frontier approach lies in i) decomposing the disturbance term into two random components representing the "random noise" and the "inefficiency" and ii) associating the frontier property with the stochastic frontier rather than the deterministic frontier. While the decomposition enables one to separate out the effects of random noise from the inefficiency and makes the support of the distribution of y independent of the parameter space, the concept of stochastic frontier ensures the frontier restriction on the observed outcomes.

Under the above assumptions, the simplest stochastic production frontier model can be represented as

$$y_i = f(x_i, \beta).exp(v_i)TE_i \quad (4.1)$$

where $f(x_i, \beta)$ is the deterministic frontier indexed by the unknown technological parameter vector β, y_i is the observed output, x_i is a vector of inputs, v_i is the random noise, $TE_i = exp(-u_i)$ is efficiency of the firm and u_i is one-sided (non-negative) latent random variable. The shortfall of the observed output (y_i) from the stochastic optimal outcome, given by $exp(-u_i)$, measures the technical efficiency of the firm characterized by stochastic elements that varies across firms.

Assuming $f(x_i, \beta)$ takes the log-linear Cobb-Douglas form (4.1) can be expressed as a linear function of the unknown parameters,

$$y_i = x_i^{'}\beta + v_i - u_i \quad (4.2)$$

where y_i is an appropriate known function of output and x_i is a vector of appropriate known functions of the inputs.

The model in (4.2) has two error components, often referred as "composed error" model. For statistical inference, the two-sided random noise (v_i) is assumed to be normally distributed and a number of probability distributions have been used to model the one-sided inefficiency (u_i). Moreover, random noise (u_i) and inefficiency (v_i) are assumed to be independent.

4.1 Estimation of the Cross-section Data SFM

Let $g(v_i, \theta)$ and $h(u_i, \eta)$ be respectively the densities of v_i and u_i indexed by the unknown parameter vectors θ and η and $\gamma=(\beta, \theta, \eta)'$ be the unknown parameter vector of the model.

The OLS would not be appropriate for estimating the parameter vector γ as the intercept term of the deterministic frontier and the non-zero mean of u_i are mixed together in the intercept term of the regression model which can not be separated out from the OLS estimate of the intercept in the regression equation. However, MOLS method of Richmond (1974) can be applied to estimate γ under specific probability distribution of 'u_i'. Thus it is necessary to specify the distributions of and for parametric estimation of SFM. Given the distributional assumptions regarding the error components, the parameters of the cross-section data SFM has been estimated by a spectrum of estimation methods viz. likelihood-based parametric, semi-parametric, Bayesian and Bayesian semi-parametric methods. In the parametric estimation of the SFM, however, the ML method, and their different variants like Simulated ML and EM method, played a dominant role, although COLS, MOLS, two-step OLS, GMM and IV methods have been used in specific situations.

The log-likelihood function of the model, based on the output y_i can be obtained from the joint probability density of (u_i, v_i) using the transformation $\varepsilon_i = y_i - x_i^{'}\beta$ and integrating out u_i and is then given by

$$l(\gamma | y) = \sum_{i=1}^{n} log \int_0^\infty g(y_i - x_i^{'}\beta + u_i, \theta) h(u_i, \eta) du_i, \gamma \in \Gamma \quad (4.3)$$

The expression (4.3) gives the general expression for the log-likelihood function of the model. The shape of the likelihood function and the nature of the likelihood equations, however, will crucially depend upon the choice of the probability densities of u_i and v_i.

While the density function of the random noise v_i has been the universally accepted as $N(0, \sigma_v^2)$, the choice of the density function for the inefficiency, however, has been varied. In fact, a sizeable portion of the works in the SFM literature has been devoted on the estimation of SFM under alternative density function of the inefficiency. For example, in normal-half-normal SFM, proposed by Aigner et al. (1977), $u_i \sim N^+(0, \sigma_u^2)$. The model was estimated using ML method via BHHH algorithm. Meeusen and van den Broeck (1977) proposed the normal-exponential SFM where $u_i \sim Exp(l/\theta)$. The parameters of the model were estimated by ML method. Greene (1997) later estimated the model by MOM.

Both the normal-exponential and the normal-half-normal SFM postulate that the mode of the inefficiency distribution is at zero which goes against the common perception of the distribution of inefficiency among firms. In response to this criticism, Stevenson (1980) proposed truncated normal i.e. $u_i \sim TrN\left(\mu_u, \sigma_u^2\right)$, which has a non-zero mode, as a plausible probability model for inefficiency and estimated the model by ML method. Later Battese and Coelli (1988) used Australian diary data to estimate the model by the COLS and the ML methods under two alternative assumptions regarding the parameter μ_u. They also performed the likelihood ratio test for $H_0 : \mu_u=0$ and found it significant for the dairy data.

The normal-gamma SFM where $u_i \sim G(P, \Theta)$, proposed by Beckers and Hammond (1987), has several attractive properties and nests the normal-exponential SFM. The log-likelihood function of the normal-gamma SFM is not only complicated but also involves integrals which has no closed form and can be approximated using some suitable numerical approximation. Greene (1990) estimated the model with approximating the intractable integral by the mathematical quadrature formula. Later Greene (2003) used Simulated Maximum Likelihood (SML) method to estimate the same model where the intractable integral in the log-likelihood function is estimated by MC simulation using Halton numbers.

5. Stochastic Frontier Model with Correlated Error Components

The noise-inefficiency independence is an important assumption that has been consistently maintained in most of studies on SFM. Recently, however, some researchers argue regarding the validity of this assumption. The logic behind this argument is that the noise-inefficiency correlation may arise due to the effect of noise which is beyond a firm's control on the efficiency which is under firm's control in taking economic decisions under uncertainty. For example, Pal and Sengupta (1999) who were first to relax the noise-inefficiency independence assumption in a simultaneous equation cross-section data SFM argued that the cropping decision in agriculture may depend upon on random factors like weather of the previous season. They estimated a stochastic production frontier model with correlated error components by systems approach using cross-section information on Indian agricultural firms and found significant noise-inefficiency correlation in the data. Subsequently Smith (2008) argued that this assumption should be relaxed at least for the empirical verification in absence of any economic logic or empirical evidence in support of any noise-inefficiency independence. Smith (2008) also found significant negative noise-inefficiency correlation in both cross-section as well as panel data SFMs.

Moreover, the effects of many institutional and social factors which are beyond the control of a firm but affect its productive efficiency are accounted for by the noise component of the composite error SFM and may lead to noise-inefficiency correlation. Finally, it is has been felt that in absence of any empirical evidence in support of the noise-inefficiency independence, one needs to develop an SFM with correlated error structure at least for empirical verification (Bandyopadhyay and Das, 2006). Furthermore, misspecification of the model is also another potential source of noise-inefficiency correlation. For example, let the true model be $y_i = a + bx_i + cz_i + v_i - u_i$ where z_i, is age of the firm, v_i is the noise and u_i is the inefficiency. Suppose u_i is uncorrelated with both x_i and v_i but correlated with z_i. Now, if the implemented model is $y_i = a + bx_i + w_i - u_i$, where $w_i = cz_i + v_i$, then clearly the error (w_i) and the inefficiency (u_i) are correlated in such a model. The noise-inefficiency correlation in this case arises due to the misspecification of the "true" model. These evidences which support the noise-inefficiency correlation will enable us to empirically verify the validity of this assumption.

Using the equation (4.2), a cross-section SFM with correlated error components can be written as:

$$y_i = x_i'\beta + v_i - u_i, \; i = 1,2,.....,n \qquad (5.1)$$

$$-\infty < v_i < \infty, \; 0 < u_i < \infty$$

where the latent random variables v_i and u_i, assumed to be statistically dependent, have the joint density function $f(v_i, u_i; \theta)$ indexed by the parameter vector θ.

It is necessary to develop the noise-inefficiency dependence structure through an appropriate joint density function of v_i and u_i to implement the model. In a few studies that have been carried out so far on correlated error components SFM in recent years such as Bandyopadhyay and Das (2006), Burns (2004), Pal (2004), Pal and Sengupta (1999), Smith (2008) three alternative approaches towards modeling noise-inefficiency dependence can be found which may be termed as: Distribution Specific Approach (DSA), Conditional-Marginal Approach (CMA), and Copula Approach (CA).

In DSA one assumes a suitable joint density function of the error components like the truncated bivariate normal distribution (Pal and Sengupta, 1999; Bandyopadhyay and Das, 2006). On the other hand in CMA, the joint density function of the error components can be obtained by combining the marginal distribution of the inefficiency with the conditional distribution of the random noise given the inefficiency (Pal, 2004). In CA, joint distribution function of the error components can be obtained by combining the marginal distributions of noise and inefficiency through a copula function (Burn, 2004; Smith, 2008).

5.1 Distribution Specific Approach (DSA)

Pal and Sengupta (1999) proposed the first ever SFM with the correlated error components which was based on the distribution specific approach. The model they used was a cross-section data simultaneous equation stochastic production frontier model which presented as $1n\ y_i = 1n\ \beta_0 + \sum_k \alpha_k\ 1n\ x_{ki} + v_i - u_i$

and $1n \frac{w_{ki}}{P_i} - 1n\ \beta_0 - 1n\ \alpha_k + 1n\ x_k - \sum_j \alpha_j\, 1n\ x_{ji} = u_i + z_{ki},\ i = 1,.....n.$

where y_i is the output with P_i as output price, x_i is a set of inputs with price w_i. Regarding the distributions of error components, it was assumed that i) (u_i, v_i) are jointly distributed as truncated bivariate normal with parameters $(\mu_u, 0, \sigma_u^2, \sigma_v^2, \rho)$ with u_i being truncated at zero, ii) technical inefficiency (u_i) and allocative efficiency (z_i) are assumed to be independently distributed and iii) $z = (z_1,.....z_m)' \sim MN(0, \Sigma)$. Making the transformation $Z_1 = u + v$ and $Z_2 = z + lu$ where $l = (1, 1,.....1)'$, the density function of (Z_1, Z_2) can be written as

$$f(Z_1, Z_2) = \left[(2\pi)^{(n+1)/2} \sigma_v \sigma_u |\Sigma|^{1/2} \sqrt{1-\rho^2}\, \Phi\left(-\frac{\mu_u}{\sigma_u}\right)\right]^{-1} \sigma_* \exp\left(-\frac{a_*}{2}\right) \Phi\left(-\frac{\mu_*}{\sigma_*}\right)$$

where

$$\frac{1}{\sigma_*^2} = \frac{1}{1-\rho^2}\left(\frac{1}{\sigma_u^2} + \frac{1}{\sigma_v^2} + \frac{2\rho}{\sigma_u \sigma_v}\right) + l' \Sigma^{-1} l$$

$$\mu_* = \sigma_*^2 \left[\frac{1}{1-\rho^2}\left\{\frac{Z_1}{\sigma_v^2}\left(1 + \rho\frac{\sigma_v}{\sigma_u}\right) + \frac{\mu_u}{\sigma_u^2}\left(1 + \rho\frac{\sigma_u}{\sigma_v}\right)\right\} + l' \Sigma^{-1} Z_2\right]$$

$$a_* = \left[\frac{1}{1-\rho^2}\left\{\frac{\mu_u^2}{\sigma_u^2} + \frac{Z_1^2}{\sigma_v^2} + 2\rho\frac{\mu_u}{\sigma_u}\frac{\mu_v}{\sigma_v}\right\} + Z_2' \Sigma^{-1} Z_2\right] -$$

$$\sigma_*^2 \left[\frac{1}{1-\rho^2}\left\{\frac{Z_1}{\sigma_v^2}\left(1 + \rho\frac{\sigma_v}{\sigma_u}\right) + \frac{\mu_u}{\sigma_u^2}\left(1 + \rho\frac{\sigma_u}{\sigma_v}\right)\frac{\mu_u^2}{\sigma_u^2}\right\} + l' \Sigma^{-1} Z_2\right]^2$$

The model was estimated by the ML method using S-PLUS software.

Das (2009) proposed a SFM with correlated error components for a single equation SFM in the line of Pal and Sengupta (1999) where v_i and u_i are assumed to be jointly distributed as truncated bivariate normal with parameter vector $\theta = (\mu_v, \mu_u, \sigma_v^2, \sigma_u^2, \rho, u_0)' \in \Theta$ where u_i is truncated below at u_0 an unknown non-negative point of truncation. The model was referred as truncated bivariate normal stochastic frontier model (TBN-SFM). The truncation point can be considered as the threshold level of inefficiency for the firms. Das (2009) analysed the specification and estimation of different sub-models under TBN-SFM.

The joint density function of and under this specification is given by

$$f(v_i, u;\ \theta) = C\ \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{v_i - \mu_v}{\sigma_v}\right)^2 - 2\rho\left(\frac{v_i - \mu_v}{\sigma_v}\right)\left(\frac{u_i - \mu_u}{\sigma_u}\right) + \left(\frac{u_i - \mu_u}{\sigma_u}\right)^2\right\}\right]$$

$$-\infty < v_i < \infty,\ u_0 < u_i < \infty$$

where,

C = $\left(2\pi\sqrt{1-\rho^2}\sigma_u\sigma_v\right)^{-1}\Phi^{-1}\left(-(u_0-\mu_u)/\sigma_u\right)$ and Φ is the distribution function of a standard normal variable.

Using the transformation and after simplification, the density function of y_i can be written as

$$f(y_i;\gamma)=\frac{1}{\sigma.\Phi\Big/\left(\alpha_1\sqrt{1+\alpha_2^2}\right)}\Phi\left(\alpha_1+\alpha_2\frac{y_i-x_i'\beta-\mu}{\sigma}\right)\phi\left(\frac{y_i-x_i'\beta-\mu}{\sigma}\right) \quad (5.2)$$

$$\gamma=(\beta,\mu_v,\mu_u,\sigma,\lambda,\rho,u_0)'\in\Gamma$$

where

$\sigma=\sigma_v\sqrt{1+\lambda^2+2\rho\lambda},\lambda=\sigma_u/\sigma_v$, $\alpha_1=-u_1^0\sqrt{1+\lambda^2+2\rho\lambda}\Big/\sqrt{1-\rho^2}$, $u_1^0=(u_0-\mu_u)/(\lambda\sigma_v)$, $\alpha_2=(\lambda+\rho)\Big/\sqrt{1-\rho^2}$, $\mu=\mu_v-\mu_u$

The density function of given in (5.2) does not follow a y_i standard distribution under the parametric specification of γ. However, under the transformation of parameter vector $\gamma=(\beta,\mu_v,\mu_u,\sigma,\lambda,\rho,u_0)'\to\delta=(\xi,\sigma,\alpha_1,\alpha_2)'$, y_i follows extended skew-normal (ESN) with parameters ξ, σ, α_1 and α_2 (Azzalini, 1985) where $\xi = x'\beta+\mu$. This re-parameterization, however, is not interest specific as the different measures of inefficiency cannot be expressed in terms of δ.

Under alternative parametric restrictions, the TBN-SFM nests a number of sub-models viz. i) with $u_0 = \mu_v = \mu_u = \rho = 0$, we get the normal-half-normal SFM (Aigner et al., 1977), ii) with $u_0 = \mu_v = \rho = 0$, we get the normal-truncated normal SFM (Stevenson, 1980), iii) with $u_0 = \mu_v = 0$, we get TBN-SFM of Pal and Sengupta (1999) and (iv) with $u_0 = \mu_v = \mu_u = 0$, we get TBN-SFM of Bandyopadhyay and Das (2006).

The moment generating function (MGF) of derived in Das (2009) and is presented as

$$M_y(t)=\left[exp\left(t\xi+\frac{t^2\sigma^2}{2}\right)\Phi\left(\frac{\alpha_1+\alpha_2 t\sigma}{\sqrt{1+\alpha_2^2}}\right)\right]\Big/\Phi\left(\frac{\alpha_1}{\sqrt{1+\alpha_2^2}}\right)$$

Using the expression of MGF of y we get the following expressions for the first four moments and the measures of skewness and kurtosis of y.

The expressions for first four moments, skewness and kurtosis are respectively given by:

$$E(y) = \xi + d\sigma h(k)$$

$$Var(y) = \sigma^2\left[1-d^2kh(k)-d^2h^2(k)\right]$$

$$\mu_3(y)=\sigma^3d^3h(k)\left[\{h(k)+k\}^2+\{h^2(k)+kh(k)-1\}\right]$$

$$\mu_4(y)=3\sigma^4+\sigma^4\left[d^4h(k)\ \left\{k\left(3+k^2\right)+4\left(1-k^2\right)h(k)-3h^2(k)\ \left(2k+h(k)\right)\right\}+6d^2h^2(k)\right]$$

$$\gamma_1(y)=\frac{d^3h(k)\left[\{h(k)+k\}^2+\{h^2(k)+kh(k)-1\}\right]}{\left[1-d^2kh(k)-d^2h^2(k)\right]^{3/2}}$$

$$\gamma_2(y)=\frac{3+\left[d^4h(k)\ \left\{k\left(3+k^2\right)+4\left(1-k^2\right)h(k)-3\ \left(2k+h(k)\right)\ h^2(k)\right\}+6d^2h^2(k)\right]}{\left[1-d^2kh(k)-d^2h^2(k)\right]^2}-3$$

where

$d=\alpha_2\Big/\sqrt{1+\alpha_2^2}$, $k=\alpha_1\Big/\sqrt{1+\alpha_2^2}$ and $h(k)$ is the hazard rate of a standard normal variate.

Putting different parametric restrictions in above equations one can get the expressions for the first four central moments, skewness and kurtosis of y for different sub-models.

Das (2009) examined analytically the identification status of the TBN-SFM and standard normal-half-normal and normal-truncated normal SFM which are nested by the TBN-SFM. It was found that i) the normal-half-normal SFM is globally identifiable, ii) normal-truncated normal SFM is locally near-identifiable and iii) TBN-SFM is either unidentifiable or near-identifiable even in a restricted parameter space. Pal and Sengupta (1999) model with restriction $u_0 = \mu_v = 0$ is near identifiable and Bandyopadhyay and Das (2006) model with restriction $u_0 = \mu_v = \mu_u = 0$ is unidentifiable.

Of interest in the SFM is the technical efficiency which is based on the conditional distribution of inefficiency given the observational error. The measure of technical efficiency due to Battese and Coelli (1988) is given by:

$$E\left(exp\left(-u_i\right)\mid\varepsilon_i\right)=\frac{1-\Phi\left(z_i+\sigma_*\right)}{1-\Phi\left(z_i\right)}exp\left(-\mu_{*i}+\frac{1}{2}\sigma_*^2\right)$$

where

$$\mu_{*i}=\frac{\mu_u\left(1-\rho\lambda\right)-\left(\varepsilon_i-\mu_v\right)\lambda\left(\lambda-\rho\right)}{\left(1+\lambda^2-2\rho\lambda\right)}$$

$$\sigma_*=\frac{\sqrt{1-\rho^2}\lambda\sigma_v}{\sqrt{1+\lambda^2-2\rho\lambda}}$$

$$z_i=\left(u_0-\mu_{*i}\right)/\sigma_*$$

and

$h(z) = \phi(z)/\Phi(-z)$ is the hazard (or failure rate) rate for a standard normal variate.

5.1.1 Estimation of Different TBN-SFM

Das (2009) proposed the EM method to estimate the parameters of the near-identifiable TBN-SFM Pal and Sengupta (1999) model. The TBN-SFM can be readily recast in "missing data" framework as the latent random variable inefficiency "u" can be considered as the variable with missing observations. Then, $y = (y_1,\ldots,y_n)'$, $u = (u_1,\ldots,u_n)'$ and $w = (y', u')'$ respectively be the observed data, the missing data and the complete data vectors and let $f(w;\theta)$, $g(y;\theta)$ and $h(u\mid y,\theta)$ be the density function of the complete data, observed data and missing data given the observed data respectively where $\theta=\left(\beta,\mu_u,\sigma_v^2,\sigma_u^2,\rho\right)\in\Theta$. Also let $l(\theta\mid w)$, $l(\theta\mid y)$ and $l(\theta\mid u, y)$ be the associated log-likelihood functions.

The proposed EM algorithm based on Dempster et al. (1977) can be described as follows:

E-Step: Let $\hat{\theta}_n$ be the estimate of θ at the nth step of iteration. Then, given $\hat{\theta}_n$, the Q-function of the model is given by

$$Q\left(\theta\mid\hat{\theta}_n\right)=\int l\left(\theta\mid w\right)h\left(u\mid y;\hat{\theta}_n\right)du$$

Note that, conditional on y, w and $\hat{\theta}_n$, the Q-function $Q\left(\theta\mid\hat{\theta}_n\right)$ is a function of θ and $\hat{\theta}_n$. M-Step: In the M-step, the Q-function is maximized with respect to θ to obtain $\hat{\theta}_{n+1}$, the (n+1)th step estimate of θ. Thus

$$\hat{\theta}_{n+1}=\arg\max Q\left(\theta\mid\hat{\theta}_n\right)$$

Given $\hat{\theta}_n$, the first order conditions for maximization of $Q\left(\theta\mid\hat{\theta}_n\right)$ are highly non-linear in $\hat{\theta}_{n+1}$, and is solved by using BHHH or BFGS algorithm. The Q-function of the model is given by (Details are found in Das (2009))

$$Q\left(\theta\mid\hat{\theta}_n\right)=-\frac{n}{2}log\left(\sigma_u^2\sigma_v^2\left(1-\rho^2\right)\right)-n\,log\,\Phi\left(\mu_u/\sigma_u\right)-\frac{1}{2\left(1-\rho^2\right)}$$

$$\left[\frac{1}{\sigma_v^2}\sum_i\varepsilon_i^2+\frac{n\mu_u^2}{\sigma_u^2}-2\mu_u\frac{\sigma_v+\rho\sigma_u}{\sigma_u^2\sigma_v}\sum_i E_{u\mid y}\left(u_i\right)\right.$$

$$-2\frac{\sigma_u+\rho\sigma_v}{\sigma_v^2\sigma_u}\sum_i\varepsilon_i E_{u\mid y}\left(u_i^2\right)+2\rho\frac{\mu_u}{\sigma_u\sigma_v}$$

$$\left.\sum_i\varepsilon_i+\frac{\sigma^2}{\sigma_u^2\sigma_v^2}\sum_i E_{u\mid y}\left(u_i^2\right)\right]$$

where $u\mid y\sim N^+\left(\mu_*,\sigma_*^2\right)$ $E_{u\mid y}\left(u_i\right)=\sigma_*\left(h\left(z\right)-z\right)$, $E_{u\mid y}\left(u_i^2\right)=\sigma_*^2\left(z^2+3zh\left(z\right)-1\right)$, $\mu_*=\left[\mu_u\sigma_v\left(\sigma_v+\rho\sigma_u\right)+\varepsilon_i\sigma_u\left(\sigma_u+\rho\sigma_v\right)\right]/\sigma^2$, $\sigma_*^2=\left[\left(1-\rho^2\right)\sigma_u^2\sigma_v^2\right]/\sigma^2$, $z=-\mu_*/\sigma_*$, and $h(z)$ is the hazard rate of a standard normal deviate.

The Bandyopadhyay & Das (2006) model can be estimated using the type-II maximum likelihood method (TML) of Berger et al. (1999) where the parameters of interest are first estimated by the integrated likelihood method and the estimated value of these parameters are used in the full likelihood function of the model to estimate the rest parameters. Das (2009) applied this method to estimate the Bandyopadhyay & Das (2006) model where the model becomes identifiable if either of the unidentifiable parameters ρ and λ is known. This feature helps to use this method to estimate the unidentified model.

The estimation method is applied treating the unidentifiable parameter λ as the nuisance parameter and the remaining parameters β, σ and ρ as the parameter of interest. The parameter vector γ can be written as $\gamma = (\gamma_1, \lambda)'$ where $\gamma_1 = (\beta, \sigma, \rho)$. Let ρ be the likelihood function of the model and $\omega(\gamma)$ be a non-negative weight function of the unidentifiable parameter λ. Then the integrated log-likelihood function (ILF) of the model is given by

$$\tilde{l}(\gamma_1) = log \int_0^\infty L(\gamma_1, \lambda)\omega(\lambda)d\lambda$$

In the first step of the TML, the identifiable parameter γ_1 is estimated by maximizing the integrated log-likelihood function and the TML estimate of γ_1 is given by

$$\hat{\gamma}_{1TML} = arg \max_{\gamma_1} \tilde{l}(y_1)$$

In the second step, $\hat{\gamma}_{1TML}$ is substituted in the log-likelihood function of the model and the unidentifiable parameter λ is estimated by maximizing the resulting log-likelihood function of the model given $\hat{\gamma}_{1TML}$ and the TML estimate of λ is given by

$$\hat{\lambda}_{TML} = arg \max_{\lambda} l(\hat{\gamma}_{1TML}, \lambda) = arg \max_{\lambda} log L(\hat{\gamma}_{1TML}, \lambda)$$

In order to implement the TML method, one needs to specify $\omega(\lambda)$, the weight function of the nuisance parameter λ. Das (2009) used one informative and one non-informative weight function for this model. The informative weight function is given by the type-2 beta density whereas the improper non-informative weight function is proportional to the reciprocal of the parameter given by λ. Das (2009) also studied the asymptotic behaviour of these estimators empirically through MC simulation and the results showed fairly good sampling properties of the estimators for moderately large samples.

The ILF under the non-information weight function was derived in Das (2009) and is given by

$$\tilde{l}(\gamma_1)\alpha - \frac{n}{2} log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2 + \sum_{i=1}^{n} log \int_0^\infty \frac{1}{\lambda}\Phi\left(\alpha_2 \frac{\varepsilon_i}{\sigma}\right)d\lambda$$

where $\varepsilon_i = y_i - x_i'\beta,\ \alpha_2 = (\lambda + \rho)\Big/\sqrt{1-\rho^2}$.

The above ILF under non-informative weights involves intractable integral and can be evaluated using the MC simulation method as it can be expressed as a expectation of some function of the parameter λ. Thus the ILF can be written as

$$\tilde{l}_2(\gamma_1)\alpha - \frac{n}{2} log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2 + \sum_{i=1}^{n} log E_\lambda\left[\Phi(\alpha_2\varepsilon_i/\sigma)/\{\lambda h(\lambda)\}\right]$$

where, $h(\lambda)$, the importance sampling density, given by the density function of $N^+(0, 1)$.

Das (2009) used the MC simulation method and the simulated ILF can be presented as

$$\hat{l}_2^s(\gamma_1)\ \alpha - \frac{n}{2} log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2 + \sum_{i=1}^{n} log \frac{1}{R}\sum_{k=1}^{R}\Phi(\alpha_{2k}\ \varepsilon_i/\sigma)/\{\lambda_k h(\lambda_k)\}$$

where $\alpha_{2k} = (\lambda_k + \rho)\Big/\sqrt{1-\rho^2}$, λ_k, $k = 1,....,R$ is a random sample of size R from $N^+(0, 1)$.

5.2 Conditional Marginal Approach (CMA)

In the conditional-marginal approach proposed in Pal (2004) the joint density function of error components are generated by combining the marginal density of inefficiency with the conditional density of the noise given inefficiency. This is much more generalised approach over distribution specific approach as different non-normal distributional assumption on the error term can be taken and joint density function can be developed through conditional distribution approach to introduce non-independence. For example, Pal (2004) developed a SFM with correlated noise-inefficiency by taking the conditional density of noise given inefficiency as normal whereas the marginal density of inefficiency as half-normal. In his specification the conditional density function of noise given inefficiency can be taken as normal with mean $\{\rho\sigma_v(u-\mu_u)\}/\sigma_u$ and variance $\sigma_v^2(1-\rho^2)$ and $u \sim N^+(0, \sigma_u^2)$ as one have bivariate normal distribution as joint density function with being the correlation coefficient. Pal (2004) also proposed a gamma distribution as the marginal distribution of u i.e. $u \sim G(P, \Theta)$ to ensure the uni-modality of the distribution which fits the situation of only a very few observations near the frontier curve and $v \mid u \sim N\left(\left\{\rho\sigma_v\left(u-\mu_u\right)\right\}/\sigma_u, \sigma_v^2\left(1-\rho^2\right)\right)$. However, they did not derive the joint distribution and marginal distribution of error under these assumptions. We described the model with the assumption of $\mu_u = 0$. The joint distribution can be obtained as follows:

$$f(u_i, v_i) = f(v_i \mid u_i) f(u_i) = \frac{\Theta^P}{\Gamma(P)\sigma_v\sqrt{1-\rho^2}} \Phi\left(\frac{v_i - \rho\sigma_v u_i/\sigma_u}{\sigma_v\sqrt{1-\rho^2}}\right) exp\left(-\theta u_i\right) u_i^{P-1}$$

Using the transformation and integrating out *ui*, the marginal density of *εi* will be

$$f(\varepsilon_i) = \frac{\Theta^P}{\Gamma(P)\sigma_v\sqrt{1-\rho^2}} \int_0^\infty \phi\left(\frac{\varepsilon_i - u_i - \rho\sigma_v u_i / \sigma_u}{\sigma_v\sqrt{1-\rho^2}}\right) exp\left(-\theta u_i\right) u_i^{P-1} du_i$$

Using the transformation $\varepsilon_i = y_i - x_i{}'\beta$, and after some simplification the density function of *yi* becomes

$$f(y_i) = \frac{\Theta^P}{\Gamma(P)\sigma_*(1-\psi)} exp\left(-\frac{1}{(1-\psi)^2}\left(\varepsilon_i\Theta - \frac{\left(1-\rho^2\right)\sigma_v^2\Theta^2}{2}\right)\right) \int_0^\infty \phi\left(\frac{u_i - \mu_{*i}}{\sigma_*}\right) u_i^{P-1} du_i$$

where

$$\varepsilon_i = y_i - x_i{}'\beta$$

$$\psi = \rho\sigma_v / \sigma_u$$

$$\mu_{*i} = \left[\varepsilon_i(1-\psi) + \left(1-\rho^2\right)\sigma_v^2\Theta\right] \Big/ (1-\psi)^2$$

$$\sigma_*^2 = \frac{\sigma_v^2\left(1-\rho^2\right)}{(1-\psi)^2}$$

The log-likelihood function for one observation is then

$$l(\gamma \mid y_i) = P\,log\,\Theta + log\,\Gamma P - log\left(\sigma_*(1+\psi)\right) +$$

$$\frac{1}{(1-\psi)^2}\left(\varepsilon_i\Theta + \frac{\left(1-\rho^2\right)\sigma_v^2\Theta^2}{2}\right) + log\int_0^\infty \phi\left(\frac{u_i - \mu_{*i}}{\sigma_*}\right) u_i^{P-1} du_i$$

where $\gamma = \left(\beta, \sigma_v^2, \sigma_u^2, P, \Theta, \rho\right)$.

This intractable integral involved in the log-likelihood function can be expressed as the expectation of a function of the random variable u_i where $u_i \sim N^+\left(\mu_{*i}, \sigma_*^2\right)$ as

$$l(\gamma \mid y_i) = P\,log\,\Theta + log\,\Gamma P - log(1+\psi) +$$

$$\frac{1}{(1+\psi)^2}\left(\varepsilon_i\Theta + \frac{\left(1-\rho^2\right)\sigma_v^2\Theta^2}{2}\right) + log\,\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right) + log\,E\left(u_i^{P-1}\right) \quad (5.3)$$

One can randomly draw a sample from this distribution and then use the simulated draws to estimate the expectation of the function with sample mean. The resulting log-likelihood function is known as the simulated log-likelihood function and given by

$$\tilde{l}(\gamma \mid y) = P\,log\,\Theta + log\,\Gamma P - log(1+\psi) +$$

$$\frac{1}{(1-\psi)^2}\left(\varepsilon_i\Theta + \frac{\left(1-\rho^2\right)\sigma_v^2\Theta^2}{2}\right) + log\,\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right) + log\frac{1}{R}\sum_{r=1}^{R} u_{ir}^{P-1} \quad (5.4)$$

where u_{ir} is a random sample of size *R* from the distribution of u_i.

Similarly, an exponential distribution as the marginal distribution of u_i can be taken by assuming $P = 1$ in the gamma model i.e. $u_i \sim Exp(\Theta)$. The corresponding log-likelihood function is then has following form

$$l(\gamma \mid y_i) = log\,\Theta - log(1+\psi) + \frac{1}{(1-\psi)^2}\left(\varepsilon_i\Theta + \frac{\left(1-\rho^2\right)\sigma_v^2\Theta^2}{2}\right) + log\,\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)$$

The simulated log-likelihood function in (5.4) is smooth continuous function of the parameter vector γ and can be maximized by using DFP or BFGS method. By the Lindeberg-Lévy variant of the Central Limit Theorem, (5.3) will be a consistent estimator of (5.4). However, as the log transformation is non-linear, the simulated log likelihood is a biased estimator of the log likelihood. As a result, the estimates obtained from maximising the simulated log likelihood are affected by simulation bias which can be minimised by increasing *R*. Moreover, the simulated maximum likelihood estimators are consistent and asymptotic normal as $n \to \infty$ and $R \to \infty$ with $\sqrt{n}/R \to 0$ (Lee, 1999). Furthermore, Bhat (2001) found that the computation time was approximately one tenth associated with using 100 Halton draws than with 1000 pseudo-random draws.

The Battese & Coelli (1988) estimator of technical efficiency, for the SFM with correlated error components is based on the conditional distribution as follows:

$$TE_i = E\left[exp\left(-u_i\right) \mid \varepsilon_i\right] = \frac{1}{f\left(\varepsilon_i\right)}\int_0^\infty exp\left(-u_i\right) f\left(u_i,\varepsilon_i\right) du_i = \frac{E\left[exp\left(-u_i\right)u_i^{P-1}\right]}{E\left[u_i^{P-1}\right]} \quad (5.5)$$

The simulated estimator of (5.5) for the ith firm is given by:

$$T\hat{E}_i = \frac{\sum_{r=1}^{R} exp\left(-u_{ir}\right)u_{ir}^{P-1}}{\sum_{r=1}^{R} u_{ir}^{P-1}}$$

5.3 Copula Approach

Smith (2008) introduced the SFM with dependent error components where the joint probability distribution of two error components is built by the copula approach of statistical modelling. In this approach the multivariate distribution function can be built using a copula function and marginal density functions of the random variables. The copula function represents the dependence structure among these random variables. Formally, an 2-variate copula function, $C_\rho\left(u_1,u_2\right)$ is a bivariate distribution function of the uniform [0,1] random variables:

$$C_\rho\left(u_1,u_2\right): \left[0,1\right]^2 \to \left[0,1\right]$$

where $u_i \sim U\left(0,1\right)$ for $i = 1,2$ and $\rho \in \Omega$. This modelling technique is based on a representation theorem of Sklar (1973). This theorem states that for every bivariate distribution function has a unique copula function that captures the dependence structure between the random variables. The joint distribution function of a set of random variables can be uniquely expressed as a function of copula function whose arguments are the marginal distribution functions of these random variables.

Let $F(u_i, v_i, \theta)$ be the joint distribution function of the random variables u_i and v_i. Then, by Sklar's representation theorem, there is a unique copula function, $C_\rho\left(u_1,u_2\right)$ so that

$$F(u_i, v_i; \theta) = C_\rho\left(H\left(u_i;\delta_1\right),G\left(v_i;\delta_2\right)\right) \quad (5.6)$$

where $H\left(u_i;\delta_1\right)$ and $G\left(v_i;\delta_2\right)$ are distribution functions of u_i and v_i respectively, ρ is the parameter that captures the dependence between u_i and v_i, $\delta = (\delta_1, \delta_2)'$ and $\theta = (\delta, \rho)'$.

The copula approach of statistical modelling uses the relation (5.3) to generate a bivariate distribution function $F(u_i, v_i; \theta)$ from a given set of distribution functions, $F(u_i, \delta_1)$ and $G(v_i, \delta_2)$ and an bivariate copula function, $C_\rho\left(u_1,u_2\right)$. The corresponding bivariate density function, $F(u_i, v_i; \theta)$ is obtained by differentiating (5.6) with respect to u_i and v_i,

$$f\left(u_i,v_i,\theta\right) = C_\rho\left(H\left(u_i;\delta_1\right),G\left(v_i;\delta_2\right)\right)h\left(u_i;\delta_1\right)g\left(v_i;\delta_2\right) \quad (5.7)$$

$$\theta = \left(\delta_1,\delta_2,\rho\right), -\infty < v_i < \infty, \ 0<u_i < \infty$$

where $h\left(u_i,\delta_1\right)$ and $g\left(v_i;\delta_2\right)$ are density functions of u_i and v_i respectively and $C_\rho\left(u_1,u_2\right) = \partial^2 C_\rho / \partial u_1 \partial u_2$ is the bivariate copula density function.

After some simplification and integrating u_i, we get the density function of y_i from (5.7) as

$$f\left(y_i,\gamma\right) = \int_0^\infty c_\rho\left(H\left(u_i;\delta_1\right),G\left(y_i - x_i'\beta + u_i,\delta_2\right)\right)h\left(u_i;\delta_1\right)g\left(y_i - x_i'\beta + u_i,\delta_2\right)du_i \quad (5.8)$$

$$\gamma = \left(\beta,\delta_1,\delta_2,\rho\right)$$

Subsequently the likelihood function of the copula-based SFM for $y_i,.....y_n$ is:

$$L\left(\gamma\right) = \sum_{i=1}^{n}\int_0^\infty c_\rho\left(H\left(u_i;\delta_1\right),G\left(y_i - x_i'\beta + u_i,\delta_2\right)\right)h\left(u_i;\delta_1\right)g\left(y_i - x_i'\beta + u_i,\delta_2\right)du_i \quad (5.9)$$

Some of the well known bivariate families of copula which can be used for modelling noise-inefficiency dependence structure are Fairlie-Gumble-Morgenstern (FGM), Ali-Mikhail-Haq (AMH), Normal, Frank, Plackett family (see: Nelson, 1999; Smith, 2008).

Smith (2008) developed the analytical expression for the density function of observational error with the assumptions that $u_i \sim Exp(\sigma_u)$ and $v_i \sim L(\sigma_v)$ under FGM copula where

$$h(u_i;\sigma_u) = \frac{1}{\sigma_u} exp\left(-\frac{u_i}{\sigma_u}\right), H(u_i;\sigma_u) = 1 - exp\left(-\frac{u_i}{\sigma_u}\right),$$

$$g(v_i;\sigma_v) = \frac{1}{\sigma_v} exp\left(-\frac{v_i}{\sigma_v}\right)\left(1 + exp\left(-\frac{v_i}{\sigma_v}\right)\right), G(v_i;\sigma_v) = \left(1 + exp\left(-\frac{v_i}{\sigma_v}\right)\right)^{-1}$$

The density function of y_i is given by

$$f(y_i,\gamma) = z_i\left[\frac{1-\rho}{\sigma_u+\sigma_v}\,{}_2F_1\left(2,1+\frac{\sigma_v}{\sigma_u};2+\frac{\sigma_v}{\sigma_u};-z_i\right)+\right.$$

$$\frac{2\rho}{\sigma_u+2\sigma_v}\,{}_2F_1\left(2,1+\frac{2\sigma_v}{\sigma_u};2+\frac{2\sigma_v}{\sigma_u};-z_i\right)+\frac{2\rho}{\sigma_u+\sigma_v}.$$

$$\left.{}_2F_1\left(3,1+\frac{\sigma_v}{\sigma_u};2+\frac{\sigma_v}{\sigma_u};-z_i\right)-\frac{4\rho}{\sigma_u+2\sigma_v}\,{}_2F_1\left(3,1+\frac{2\sigma_v}{\sigma_u};2+\frac{2\sigma_v}{\sigma_u};-z_i\right)\right]$$

where

$$z_i = exp\left(\frac{y - x_i'\beta}{\sigma_v}\right)$$

and

$${}_2F_i(a,b;;c;s) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-st)^{-a}\,dt = \sum_{i=0}^{\infty}\frac{(a)_i(b)_i}{(c)_i}\frac{s^i}{i!}$$

One difficulty with the copula-based SFM, however, is that the log-likelihood function based on the density function of y_i of the model involves intractable integrals which are to be evaluated either numerical method or MC method. Smith (2008) estimated the model by ML method approximating the log-likelihood function using numerical integration. The ML method was performed using BFGS algorithm. The standard normal-half-normal distribution was used for the distributions of the error components in a cross section study of US electricity data with the AMH, the Frank family and the Plackett families of copula.

However, this method can be computationally burdensome in case of certain copula functions. As an alternative, simulation estimation procedure can be used. Simulated maximum likelihood (SML) is one of the popular methods which uses simulation technique to approximate an integral arises in the likelihood function and have no closed form. The estimation procedure involves transforming the integral as an expectation with respect to the distribution of a random variable. One can randomly draw from this distribution and then use the simulated draws to estimate the expectation with a sample mean. The log-likelihood function of (5.6) can be written as

$$L(\gamma) = \sum_{i=1}^{n} E_{ui}\left[c_\rho\left(H(u_i;\delta_1),G(y_i - x_i'\beta + u_i,\delta_2)\right)g(y_i - x_i'\beta + u_i,\delta_2)\right] \quad (5.10)$$

where the $E_{ui}[.]$ is the expectation with respect to the distribution of u_i.

Burns (2004) used Halton sequence based SML technique for estimating the parameters copula-based cross section data SFM. The simulated likelihood function of (5.8) can be written as

$$L(\gamma) = \sum_{i=1}^{n}\frac{1}{R}\sum_{k=1} c_\rho\left(H(u_{ik};\delta_1),G(y_i - x_i'\beta + u_{ik},\delta_2)\right)g(y_i - x_i'\beta + u_{ik},\delta_2)$$

where u_{ik} are random draws from the distribution of u_i and R is the number of random draws used in the estimation. Burn (2004) used Halton sequence (Halton, 1960) based random draws which provides computational efficiency and accuracy than Uniform random draws in terms of simulation. Bhat (2001) found more accurate result with 100 Halton draws than 1000 pseudo-random draws. Train (1999) and Sandor and Train (2004) confirmed this result in the context of multinomial logit models. Burn used the model of Smith (2008) to illustrate the proposed SML method to estimate SFM with correlated error components using US electricity data. The model was estimated under AMH, Plackett, Normal, FGM and Frank families of copula using normal-half-normal error components.

Burns (2004) developed the analytical expression for the density function of y_i with the assumptions that $u_i \sim N^+\left(0,\sigma_u^2\right)$ and $v_i \sim N\left(0,\sigma_v^2\right)$ under FGM copula where

$$h\left(u_i;\sigma_u^2\right)=\frac{2}{\sigma_u}\phi\left(\frac{u_i}{\sigma_u}\right),H\left(u_i;\sigma_u^2\right)=2\Phi\left(\frac{u_i}{\sigma_u}\right)-1,g\left(v_i;\sigma_v^2\right)=\frac{1}{\sigma_v}\phi\left(\frac{v_i}{\sigma_v}\right),\ G\left(v_i;\sigma_v^2\right)=\Phi\left(\frac{v_i}{\sigma_v}\right)$$

The density function of y_i is given by

$$f\left(y_i;\gamma\right)=E_u\left[\phi\left(\frac{y_i-x_i'\beta+u_i}{\sigma_v}\right)\left\{1+\rho\left(1-2\Phi\left(\frac{y_i-x_i'\beta+u_i}{\sigma_v}\right)\right)\left(3-4\Phi\left(\frac{u_i}{\sigma_u}\right)\right)\right\}\right]$$

The corresponding simulated likelihood function can be written as

$$\tilde{L}(\gamma)=\sum_{i=1}^{n}\frac{1}{R}\sum_{k=1}\left[\phi\left(\frac{y_i-x_i'\beta+u_{ik}}{\sigma_v}\right)\left\{1+\rho\left(1-2\Phi\left(\frac{y_i-x_i'\beta+u_{ik}}{\sigma_v}\right)\right)\left(3-4\Phi\left(\frac{u_{ik}}{\sigma_u}\right)\right)\right\}\right]$$

The Battese & Coelli (1988) estimator of technical efficiency, for the SFM with correlated error components under copula approach can be expressed as follows:

$$TE_i=E\left[exp\left(-u_i\right)\mid\varepsilon_i\right]=\frac{1}{f\left(\varepsilon_i\right)}\int_0^{\infty}exp\left(-u_i\right)f\left(u_i,\varepsilon_i\right)du_i=$$

$$\frac{E_{ui}\left[exp\left(-u_i\right)c_\rho\left(H\left(u_i;\delta_1\right),G\left(\varepsilon_i+u_i,\delta_2\right)\right)g\left(\varepsilon_i+u_i,\delta_2\right)\right]}{E_{ui}\left[c_\rho\left(H\left(u_i;\delta_1\right),G\left(\varepsilon_i+u_i,\delta_2\right)\right)g\left(\varepsilon_i+u_i,\delta_2\right)\right]} \qquad (5.11)$$

The simulated estimator of (5.11) for the ith firm is given by:

$$T\hat{E}_i=\frac{\sum_{r=1}^{R}\left[exp\left(-u_i\right)c_\rho\left(H\left(u_i;\delta_1\right),G\left(\varepsilon_i+u_i,\delta_2\right)\right)g\left(\varepsilon_i+u_i,\delta_2\right)\right]}{\sum_{r=1}^{r}\left[c_\rho\left(H\left(u_i;\delta_1\right),G\left(\varepsilon_i+u_i,\delta_2\right)\right)g\left(\varepsilon_i+u_i,\delta_2\right)\right]}$$

6 Empirical Evidence

This section presents an empirical application of the three different approaches of SFM with correlated error components discussed above using much analysed cross-section information on the 123 US electricity firms given in Greene (1990). The stochastic cost frontier model (Greene 1990, pp. 154) estimated is given by

$$In\left(Cost/P_f\right)=\beta_0+\beta_1\ 1n\left(Q\right)+\beta_2\ 1n^2\left(Q\right)+\beta_3\ 1n\left(P_l/P_f\right)+\beta_4\ 1n\left(P_k/P_f\right)+u+v$$

where Q is output and P_l, P_k and P_f are prices of three factor inputs labor *(l)*, capital *(k)* and fuel *(f)*. It is assumed that the noise component *(v)* is distributed as N $\left(0,\sigma_v^2\right)$ and the technical efficiency component *(u)* is distributed as $N^+\left(0,\sigma_u^2\right)$. Also, a parametric transformation from $\left(\sigma_u^2,\sigma_v^2\right)\rightarrow\left(\sigma,\lambda\right)$ where $\sigma = \sigma_v\sqrt{1+\lambda^2+2\rho\lambda}$ and $\lambda=\sigma_u/\sigma_v$ is used for the estimation of model under different approaches. (Table 1)

The estimates of the parameters were derived using non-information weight under DSA approach. The model

Table 1: Estimates of Stochastic Cost Frontier of the US Electricity Industry*

Parameter	Estimate of the Parameters			
	Product	DSA	CMA	Copula
β_0	-7.391 (0.293)	-7.413 (0.327)	-7.424 (0.329)	-7.524 (0.316)
β_1	0.405 (0.028)	0.427 (0.031)	0.422 (0.034)	0.418 (0.032)
β_2	0.031 (0.003)	0.030 (0.002)	0.031 (0.002)	0.032 (0.003)
β_3	0.244 (0.065)	0.248 (0.027)	0.244 (0.033)	0.254 (0.043)
β_4	0.061 (0.068)	0.048 (0.049)	0.059 (0.049)	0.062 (0.065)
σ	0.186 (0.029)	0.215 (0.038)	0.209 (0.042)	0.205 (0.057)
λ	1.350 (0.148)	0.905 (0.134)	1.114 (0.174)	1.223 (0.192)
ρ	0	-0.326 (0.157)	-0.366 (0.177)	-0.393 (0.195)
Log-L	67.163	69.276	69.878	70.966

**Figures in the bracket denote the estimated standard errors* was also estimated under CMA and Copula approaches using the SML method. The bivariate family of FGM copula is chosen for this analysis. The ML estimates of parameters under different approaches, along with their asymptotic standard errors, are reported in Table-1. It can be noted that the first column presents estimates of parameters under the standard SFM with independent error components and next three columns present estimates of parameters under SFM with three different approaches of noise-inefficiency correlation. The result shows that magnitudes of the estimates of the slope coefficients do not differ significantly for each SFM with three different approaches of noise-inefficiency correlation. Also the variance of different SFM with correlated error components have higher values compare to the variance of standard SFM due to correlation structure. The estimates of correlation parameter are consistent with negative dependence between the error components for these data, although the estimated standard error on each is obviously fairly large. Moreover, the presence of the noise-inefficiency correlation have substantial effect on the estimates of the parameters of the distribution of the error components; this effect being most pronounced for the parameter. Table 2 provides descriptive statistics of estimates of cost efficiency for all 123 firms. This shows that the correlated error components have an effect on cost efficiency estimates which is evident from the declining sample means.

7 Conclusions

This paper presents the early developments of the SFM. It discusses the development of SFM with correlated error components developed so far. The modelling of SFM with correlated error components can be classified in three different approaches. Among these approaches DSA assumes a suitable joint distribution function of the error components like the truncated bivariate normal distribution (Bandyopadhyay and Das, 2006; Pal and Sengupta, 1999). In CMA the joint distribution function of the error components can be obtained by combining the marginal distribution of the inefficiency with the conditional distribution of the random noise given the inefficiency (Pal, 2004). In CA, joint distribution function of the error components can be obtained by combining the marginal distributions of noise and inefficiency through a copula function (Smith, 2008). The copula approach is the most attractive one as a variety of different models can be derived changing the marginal distribution of inefficiency and/or copula function.

In addition to further developments of this model, there are a number of other areas of potential research. These include extending the model with an upper bound to inefficiency or a lower bound to the efficiency in a cross-sectional correlated error components context as proposed by Almanidis et al. (2011) for stochastic frontier model with independent error components. Moreover, in these limited study of correlated error components SFM, heteroskedasticity in error components were not tried out. The heteroskedasticity problem is severe in SFM and there have been a number of studies taking heteroskedastic inefficiency in standard SFM structure. This can be extended to correlated error components SFM. The distribution of composite error become more complex in nature under SFM with correlated error components and different estimation procedures should be carried out for comparison.

Table 2: Descriptive Statistics of Estimate of Cost Efficiency

Selected Features	Cost Efficiency			
	Product	DSA	CMA	Copula
Mean	0.8579	0.8349	0.8468	0.8435
Std. Dev.	0.0656	0.0696	0.0687	0.0731
Minimum	0.6532	0.6279	0.5816	0.5622
Maximum	0.9844	0.9653	0.9548	0.9523

References

Aigner, D. and S. F. Chu (1968). On estimating the industry production function. *American Economic Review,* 58(4), 826-839.

Aigner, D. C. A. K. Lovell and P. Schmidt (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics,* 6, 21-37.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics,* 12, 171-178.

Bandyopadhyay, D. and A. Das (2006). On measures of technical inefficiency and production uncertainty in stochastic frontier production function with correlated error components. *Journal of Productivity Analysis,* 26(2), 165-180.

Bandyopadhyay, D. and A. Das (2009). Estimation of the truncated bivariate normal stochastic frontier model by EM algorithm. *Journal of Quantitative Economics,* 7(2), 80-95.

Battese, G. and T. Coelli (1988). Prediction of firm level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics,* 38, 387-399.

Beckers, D. and C. Hammond (1987). A tractable likelihood function for the normal-gamma stochastic frontier model. *Economics Letters,* 24, 33-38.

Berger, O. J., B. Liseo and R. L. Wolpert (1999). Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science,* 14(1), 1-28.

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research,* 35, Part B, 677-693.

Brissimis, S. N., M. D. Delis and E. G. Tsionas (2010). Technical and allocative efficiency in European banking, *European Journal of Operation Research,* 204(1), 153-163.

Burns, R. C. J. (2004). The simulated maximum likelihood estimation of stochastic frontier models with correlated error components. Unpublished Dissertation, Department of Econometrics and Business Statistics, The University of Sydney, Australia.

Camanho, A. S., and R. G. Dyson (2005). Cost efficiency, production and value-added models in the analysis of bank branch performance. *Journal of the Operational Research Society,* 56, 483-94.

Choi, K.-S. (2011). *Estimation of managerial efficiency in baseball: A bayesian approach*. Working paper, Department of Economics, Arizona State University.

Coelli, T., D. S. P. Rao and G. E. Battese (1998). *An introduction to efficiency and productivity analysis*. Kluwer Academic Publishers, Boston.

Cummins, J. D., M. Rubio-Misas and H. Zi (2005). The Effect of Organizational Structure on Efficiency: Evidence from the Spanish Insurance Industry. *Journal of Banking & Finance*, 28(12), 3113-3150.

Dar, A. A. (2007). Estimating worker information gaps from a stochastic wage frontier: A study of canadian labour markets. International Business & Economics Research Journal, 6(12), 91-98.

Das A. (2009). Contribution to econometric analysis of stochastic frontier model with correlated error components, Ph. D. thesis, Department of Statistics, University of Kalyani, Kalyani.

Dawson, P. & S. Dobson (2002). Managerial efficiency and human capital: An application to english association football. Managerial and Decision Economics, 23, 471-486.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, 19(3), 273-292.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society*, Series B, 39, 1-38.

Dolton, P., O. D. Marcenaro, L. Navarro (2003). The effective use of student time: A stochastic frontier production function case study. *Economics of Education Review*, 22(6), 547-560.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, Series A, 120(3), 253-281.

Fried, H. O., C. A. K. Lovell, S. S. Schmidt (1993). *The measurement of productive efficiency: Technique and applications*. Oxford University Press, New York.

Greene, W. H. (1990). A gamma distributed stochastic frontier model. *Journal of Econometrics*, 46, 141-164.

Greene, W. H. (2003). Simulated likelihood estimation of the normal-gamma stochastic frontier function. *Journal of Productivity Analysis*, 19, 179-190.

Halton, J. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2, 84-90.

Jondrow, J., C. A. K. Lovell, I. Materov and P. Schmidt (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19, 233-238.

Khatri, Y., L. Leruth, and J. Piesse (2002). *Corporate performance and governance in malaysian*. IMF working paper, Asia and Pacific Department.

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities, in T. C. Koopmans, ed., Activity analysis of production and allocation, cowles commission for research in economics, Monograph No. 13, New York: Wiley.

Kumbhakar, S. C. and C. A. K. Lovell (2000). Stochastic frontier analysis. Cambridge University Press.

Lee, L.-F. (1999). Statistical inference with simulated Likelihood Functions. *Econometric Theory*, 15, 337-360.

Meeusen, W. and J. ven den Broeck. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18, 435-444.

Murillo-Zamorano, L. R. (2004). Economic efficiency and frontier techniques. *Journal of Economic Surveys*, 18, 33-77.

Nelson, R. B. (1999). *An introduction to copulas*. Springer-Verlag: New York.

Pal, M. (2004). A note on a unified approach to the frontier production function models with correlated non-normal error components: The case of cross section data. *Indian Economic Review*, 39, 7-18.

Pal, M. and A. Sengupta (1999). A model of FPF with correlated error components: An application to indian agriculture. *Sankhya*, 61, Series B, 337-350.

Richmond, J. (1974). Estimating the efficiency of production. *International Economic Review*, 15(2), 515-521.

Schmidt, P. (1976). On the statistical estimation of parametric frontier production functions. *Review of Economics and Statistics*, 58(2), 238-239.

Schmidt, P. and T. F. Lin (1984). Simple tests of alternative specifications in stochastic frontier models. *Journal of Econometrics,* 24(3), 349-361.

Shephard, R. W. (1953). *Cost and production functions*. Princeton University Press, Princeton.

Shephard, R. W. (1970). *The theory of cost and production functions*. Princeton University Press, Princeton.

Silva Portela, M. C. A., and E. Thanassoulis (2005). Profitability of a sample of portuguese bank branches and its decomposition into technical and allocative components. *European Journal of Operational Research* 162(3), 850-66.

Sklar, A. (1973). Random variables, joint distributions, and copulas. *Kybernetika*, 9, 449-460.

Smith, M. D. (2008). Stochastic frontier model with dependent error components. *Econometrics Journal*, 11, 172-192.

Stevenson, R. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, 13, 57-66.

Troutt, M. D., M. Y. Hu and M. S. Shanker (2005). A distribution-free approach to estimating best response values with application to mutual fund performance modeling. *European Journal of Operational Research* 166(2), 520-27.

Arabinda Das has been working as Assistant Professor in the Department of Statistics, Acharya Prafulla Chandra College, New Barrackpore, Kolkata - 700 131, West Bengal. He did his Ph.D. in Statistics in the area of Stochastic Frontier Model. His research interests are efficiency and productivity analysis, modeling dependence through copula function, simulation based estimation etc. His e-mail id is ara.das@gmail.com.