Stochastic Frontier Models with Correlated Error Components

by

Murray D Smith[1]

May 17, 2004

Abstract: In the productivity modelling literature, the disturbances U (representing technical inefficiency) and V (representing noise) of the composite error $W = V - U$ of the stochastic frontier model are assumed to be independent random variables. By employing the copula approach to statistical modelling, the joint behaviour of U and V can be parameterised thereby allowing the data the opportunity to determine the adequacy of the independence assumption. In this context, three examples of the copula approach are given: the first is algebraic (the Logistic-Exponential stochastic frontier model with margins bound by the Fairlie-Gumbel-Morgenstern copula) and the second and third are empirically oriented, using data sets well-known in productivity analysis. Analysed are a cross-section of cost data sampled from the US electrical power industry, and an unbalanced panel of data sampled from the US airline industry.

Keywords: Stochastic frontier model; Copula; Sklar's theorem; Copula approach; Spearman's S_ρ.

JEL Classification Codes: C21, C23, C51

[1]Address for correspondence: Econometrics and Business Statistics, The University of Sydney, Sydney NSW 2006, Australia. (E-mail: Murray.Smith@econ.usyd.edu.au). For constructive and helpful comments, thanks are due to Mark Neal, Chris O'Donnell and seminar participants at the University of Sydney. Any errors in the manuscript are my responsibility.

1 Introduction

In the productivity modelling literature - see Coelli et al [3], Kumbhakar and Knox Lovell [17] and, more recently, Murillo-Zamorano [20] and the many references in each - the disturbances U and V representing technical inefficiency and noise, respectively, of the composite error

$$W = V - U$$

of the stochastic frontier model (SFM hereafter) are assumed to be independent random variables. This ubiquitous assumption, held apparently without question in the SFM literature, is moreover described by Kumbhakar and Knox Lovell as "seems innocuous" [17, p. 75]. Despite that opinion, the lack of factual evidence to support the independence assumption in the SFM is cause for concern, and so it is the need to examine the consequences of weakening this assumption that motivates the work presented in this article. Independence is relaxed by introducing estimable parametric forms to represent association between the error components, each with special, leading cases that represent independence. In doing so, the data themselves are allowed the opportunity to determine statistically the adequacy of the independence assumption.

The joint behaviour of U and V can be parameterised by employing the copula approach to statistical modelling. This technique derives from a representation theorem due to Sklar, see [24] and [25], in which the joint distribution of random variables can be expressed as a function of its univariate margins: that function being the *copula*. The copula represents the dependence structure that associates random variables, it captures entirely their joint behaviour. Whilst there exists an extensive statistical literature on copulas, they have to date received very little attention in econometrics. Applications include Dardanoni and Lambert [5] in economics, and Miller and Liu [19] and Smith [28] in econometrics. In finance, Embrechts et al [6] use copulas to model risk. In regard to statistical modelling, Joe [14, Chapter 11] gives five studies in which copula functions are used to model various multivariate and longitudinal data sets.

The copula approach to modelling derives from the converse of Sklar's theorem, in that specifying marginal distributions for each random variable and a copula function that binds them, it follows that this process must yield a joint distribution; i.e. this method constructs a statistical model. The SFM is a prime candidate for the application of the copula approach as it is standard procedure to specify pairs of models for the error components. Common choices for U include the Exponential, the Half-Normal, the Truncated Normal and the Gamma distributions, and for V it is typically the Normal distribution. Any pairing is permitted under a copula approach, including pairs picked from the aforementioned distributions.

Three major examples are given: the first is algebraic, the remaining are data-driven. The algebraic example concerns the Logistic-Exponential SFM, whose margins are bound by the Fairlie-Gumbel-Morgenstern family of copulas. This model is convenient in that closed form expressions for the density of W and the Battese-Coelli measure of technical efficiency can be derived. Both are studied allowing for differing degrees of positive and negative association between U and V.

The data-driven examples use relatively well-known and readily available data: the US Electricity Utility cross-section, and the US Airlines panel. The cross-section example focuses on the Normal-Half-Normal SFM, with margins bound by differing copulas in order to explore model choice. An interesting outcome from this example is the difference between the kernel-smoothed distributions of technical efficiency. For the preferred model (that estimates significant negative association between the error components) the distribution is located substantially away from its standard (unassociated) Normal-Half-Normal counterpart.

The final example focuses on the Normal-Truncated Normal SFM applied to unbalanced panel data; collected on n firms across T time periods. In the context of the mutually independent random effects formulation of Kumbhakar [16] with time-varying technical efficiency, the desired structure of association between a firms error components is common association between the time invariant inefficiency error component and all time-varying noise components, with the noise component serially independent. To represent this structure it is necessary to specify a $(T+1)$-dimensional copula, the two considered are the extended Fairlie-Gumbel-Morgenstern family of copulas and the Meta-Gaussian family of copulas.

2 Copula Functions

2.1 Theory

The study of the copula function was initiated by Hoeffding in the 1940s, and further developed by Fréchet in the post-war period. Particularly important was the work of Sklar, especially his representation theorem from which the copula approach to modelling is derived; see [24] and [25]. For histories of the development of copula theory see Dall'Aglio [4], Schweizer [22], Fisher [8] and Nelsen [21]; also of interest is Sklar [26]. Joe [14], Frees and Valdez [9] and Nelsen [21] present comprehensive surveys of the theory of copula functions.

The simplest case will be set down here to illustrate the theory: the bivariate case. Consider a two-place function $C : \mathbb{I}^2 \to \mathbb{I}$, where $\mathbb{I}$ denotes the closed interval $[0,1]$ of $\mathbb{R}$, the latter denoting the real line, while $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ will later be used to denote the extended real line. C is a copula function if it is 2-increasing with margins $C(1,y) = y$ and $C(x,1) = x$, and grounded such that $C(0,y) = C(x,0) = C(0,0) = 0$, where the pair $(x,y) \in \mathbb{I}^2$. By 2-increasing it is meant

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0$$

for every x_1, x_2, y_1, y_2 in $\mathbb{I}$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$.

Sklar's main result is that there exists a copula function which acts to represent the joint cumulative distribution function (cdf hereafter) of random variables in terms of its underlying one-dimensional margins. For example, let $F_1(z_1)$ and $F_2(z_2)$ denote, respectively, the cdf of the random variables Z_1 and Z_2; that is, $F_i(z_i) = \Pr(Z_i \leq z_i)$, where $z_i \in \overline{\mathbb{R}}$ $(i = 1,2)$, and let $F(z_1, z_2) = \Pr(Z_1 \leq z_1, Z_2 \leq z_2)$ denote the joint cdf. Sklar's result is that the joint cdf can be represented according to

$$F(z_1, z_2) = C(F_1(z_1), F_2(z_2)). \tag{1}$$

The copula representation is a re-formulation of the joint cdf such that it separates the margins F_1 and F_2 from their interaction. So while the copula function takes as arguments the margins F_1 and F_2 in (1), the function itself is independent of those margins. The copula serves to capture the association between the random variables Z_1 and Z_2. When F_1 and F_2 are continuous functions, then (1) provides a unique representation of the cdf for any $(z_1, z_2) \in \overline{\mathbb{R}}^2$. Nelsen [21, Section 2.3] provides a proof of (1) that follows the method given in Schweizer and Sklar [23, Chapter 6] where a multivariate version of the theorem is proved. For an alternate proof (multivariate version) see Carley and Taylor [2].

The copula density $c : \mathbb{I}^2 \to [0, \infty]$ of a copula C is defined as

$$c(x,y) = \frac{\partial^2}{\partial x \partial y} C(x,y).$$

It cannot be negative-valued as C is 2-increasing. The copula density occurs in the expression for the joint probability density function (pdf hereafter) of continuous random variables. Assuming that F_1 and F_2 are (right) continuous functions, then, from (1), the joint pdf of Z_1 and Z_2 is given by

$$\begin{aligned} \frac{\partial^2}{\partial z_1 \partial z_2} F(z_1, z_2) &= f_1(z_1) f_2(z_2) \left. \frac{\partial^2}{\partial x \partial y} C(x,y) \right|_{x \to F_1(z_1), y \to F_2(z_2)} \\ &= f_1(z_1) f_2(z_2) c(F_1(z_1), F_2(z_2)) \end{aligned}$$

where $f_i(z_i) = \frac{\partial}{\partial z_i} F_i(z_i)$ denotes the pdf of Z_i, $i = 1, 2$. Fang et al [7] call $c(F_1(z_1), F_2(z_2))$ the "density weighting function".

2.2 Examples

Consider the Product copula:

$$\Pi = xy$$

where here, and in the following examples, $(x, y) \in \mathbb{I}^2$. In light of (1), under Π the joint cdf of Z_1 and Z_2 must be given by $F(z_1, z_2) = F_1(z_1)F_2(z_2)$, implying that Z_1 and Z_2 are independent. Thus, the Product copula represents (bivariate) independence. The copula density of Π is obviously equal to unity.

A second example illustrates the re-formulation as per (1). Let $(Z_1, Z_2) = (z_1, z_2) \in \overline{\mathbb{R}}^2$ have joint cdf

$$\begin{aligned} F(z_1, z_2) &= (1 + e^{-z_1} + e^{-z_2})^{-1} \\ &= \left. \frac{xy}{x + y - xy} \right|_{x \to F_1(z_1), y \to F_2(z_2)} \end{aligned}$$

where $F_i(z_i) = (1 + e^{-z_i})^{-1}$ $(i = 1, 2)$ are the Logistic margins of F. The copula in this case is clearly $xy/(x + y - xy)$, and it may be written

$$\frac{\Pi}{\Sigma - \Pi} \tag{2}$$

where $\Sigma = x + y$. The copula density in this case is

$$\frac{2\Pi}{(\Sigma - \Pi)^3}.$$

Two further examples are the (bivariate) Fréchet lower bound

$$W = \max(x + y - 1, 0)$$

and the (bivariate) Fréchet upper bound

$$M = \min(x, y).$$

These copulas are important in that the closed interval $[W, M]$ has the property of containing all bivariate copulas; namely, $W \leq C(x, y) \leq M$ for all $(x, y) \in \mathbb{I}^2$.

2.3 Families of Copulas

For the purposes of statistical modelling it is desirable to parameterise the copula function so that data can be used to shed light on the extent of association between the random variables of interest. Let

$$C_\theta(x, y)$$

denote a family of copulas, where the members are indexed according to values assigned to θ (possibly vector valued). Provided that the margins F_1 and F_2 do not depend on θ, Sklar's representation (1) holds for all members of a given family.

There are numerous examples of families of bivariate copulas given in Joe [14] and Nelsen [21]. For example, the family of Bivariate Normal copulas is given by

$$\Phi_2(\Phi^{-1}(x), \Phi^{-1}(y); \theta) \tag{3}$$

where $-1 \leq \theta \leq 1$. Here, $\Phi(\cdot)$ denotes the cdf of a standard Normal variable, and $\Phi_2(\cdot, \cdot; \theta)$ the cdf of a bivariate standard Normal variable with Pearson's product moment correlation coefficient θ. Note that setting $x = \Phi(z_1)$ and $y = \Phi(z_2)$ in (3) recovers the bivariate standard Normal cdf, $\Phi_2(z_1, z_2; \theta)$. The copula density of the Bivariate Normal copula is given by

$$\frac{\phi_2\left(\Phi^{-1}(x), \Phi^{-1}(y); \theta\right)}{\phi\left(\Phi^{-1}(x)\right)\phi\left(\Phi^{-1}(y)\right)}$$

where $\phi(\cdot)$ denotes the pdf of a standard Normal variate. Replacing x and y with, for example, the cdfs $F_1(z_1)$ and $F_2(z_2)$, respectively, yields the bivariate Meta-Gaussian distribution that Fang et al [7] attribute to Krzysztofowicz and Kelly.

A second example is the Ali-Mikhail-Haq (AMH hereafter) family of copulas:

$$\frac{xy}{1-\theta(1-x)(1-y)} \tag{4}$$

where the association parameter is such that $-1 \leq \theta < 1$. Setting $\theta = 0$ yields Π, and allowing $\theta \to 1$ yields (2) as the limit. The copula density of the AMH copula is

$$\frac{1+\theta\left(xy+x+y-2+\theta\left(1-x\right)\left(1-y\right)\right)}{\left[1-\theta\left(1-x\right)\left(1-y\right)\right]^3}.$$

Continuing the earlier example on the Bivariate Logistic distribution, the AMH (in this case with $-1 \leq \theta \leq 1$) is the copula of a generalised version of that distribution with cdf

$$\begin{aligned} F(z_1, z_2) &= (1+e^{-z_1}+e^{-z_2}+(1-\theta)e^{-z_1-z_2})^{-1} \\ &= \left.\frac{xy}{1-\theta(1-x)(1-y)}\right|_{x \to F_1(z_1), y \to F_2(z_2)} \end{aligned}$$

where, as was the case before, both margins F_1 and F_2 are Logistic.

2.4 Measuring Association

The ability of a given family of (bivariate) copulas to represent differing degrees of association can be examined in terms of the extent to which it covers, for every $(x, y) \in I\!I^2$, the interval between the lower and upper Fréchet bounds for copulas, $[W, M]$. This is generally determined at the extremes of the parameter space. For example, the Bivariate Normal family (3) has full coverage as $\Phi_2(\Phi^{-1}(x), \Phi^{-1}(y); -1) = W$ and $\Phi_2(\Phi^{-1}(x), \Phi^{-1}(y); 1) = M$. Furthermore, this family is comprehensive because it also includes $\Pi = \Phi_2(\Phi^{-1}(x), \Phi^{-1}(y); 0)$. Comprehensive families of copulas parameterise the full range of association. Despite the existence of comprehensive families, it may be counterproductive in modelling contexts to confine attention to such families, for there are typically many other features of the data that are of interest.

Many copula families are not comprehensive, one example being the AMH family (4): it includes Π, when $\theta = 0$, but it fails to contain either W or M. For such families it is desirable to assess coverage in terms of measures of association. In this respect, most familiar is Pearson's product moment correlation coefficient. However, this measure suffers from a lack of invariance with respect to the margins. A numerical example illustrates: suppose that the copula of the joint distribution of $Z_1 \sim \mathit{Logistic}$ and $Z_2 \sim \mathit{Exponential}(1)$ is the AMH. For this model, Pearson's measure is bounded such that $[-0.2405, 0.3556]$, to 4dp. Now alter only the Z_1 margin, and suppose that $Z_1 \sim N(0, 1)$. For this second model the corresponding region is $[-0.2459, 0.3616]$, to 4dp. This implies that when dealing with models that depart from the multivariate Normal, that it is advisable to seek alternatives to Pearson's measure that satisfy invariance.

Two measures that are invariant are Kendall's τ and Spearman's S_ρ. Both are concordance measures that are bounded to the interval $[-1, 1]$: equal to -1 at W, 1 at M and 0 for Π. Moreover, both depend only on the copula of the joint distribution. Focusing on S_ρ, for independent pairs (Z_{1i}, Z_{2i}), $i = 1, 2, 3$ that are copies of (Z_1, Z_2), it is defined as (e.g. see Nelsen [21, Section 5.1]):

$$S_\rho = 3\left[\Pr((Z_{11} - Z_{12})(Z_{21} - Z_{23}) > 0) - \Pr((Z_{11} - Z_{12})(Z_{21} - Z_{23}) < 0)\right].$$

Provided Z_1 and Z_2 are continuous random variables, with the copula of their joint distribution given by C, then it can be shown that

$$\begin{aligned} S_\rho &= 12\int_{\boldsymbol{I}^2} xy dC(x, y) - 3 \\ &= 12E[XY] - 3 \end{aligned}$$

where the second line forms S_ρ as a product moment, by interpreting X and Y as standard Uniform random variables with joint cdf C. To illustrate, for the AMH family it can be shown that

$$S_\rho = \frac{12(1+\theta)}{\theta^2}\operatorname{dilog}(1-\theta) - \frac{24(1-\theta)}{\theta^2}\log(1-\theta) - \frac{3(\theta+12)}{\theta}$$

where $\operatorname{dilog}(z) = \int_1^z \log(t)\,(1-t)^{-1}\,dt$ is the dilogarithm function. Under the AMH copula, S_ρ is bounded such that $[-0.2711, 0.4784]$, to 4dp, irrespective of the margins of the model.

2.5 The Copula Approach to Modelling

For the purposes of modelling it is the converse of the copula representation of the joint cdf given by Sklar's theorem that is relevant. In other words, given models for the margins and a copula function that binds them, this then has the effect of constructing a statistical model for the random variables of interest, as a joint cdf is specified. For example, if Z_1 and Z_2 denote the variables of interest, then the statistical model for Z_1 and Z_2 is their true, but unknown joint distribution; naturally, this distribution may depend on parameters and covariates. Under a copula approach, separate models for the margins $F_1(z_1)$ and $F_2(z_2)$ are proposed, as well as a selection of a copula family C_θ. Combining these selections as per (1) then has the effect of specifying the joint cdf of Z_1 and Z_2. Intuitively, the copula approach determines each component of the overall model, then engineers them together using a copula function.

3 The Stochastic Frontier Model

The Stochastic Frontier Model (SFM) is given by

$$\log Y = x'\beta + V - U \tag{5}$$

where the single-valued output $Y = y \in \mathbb{R}_+$; note that any unit of measurement associated with output is ignored in this model merely for notational ease. x $(k \times 1)$ is a vector of regressors (i.e. known functions of the inputs) assumed exogenous, and β $(k \times 1)$ is a vector of unknown parameters (that the regression function in (5) is linear in β is unimportant in what follows, and can be relaxed if required). In this model, the error component $U = u \in \underline{\mathbb{R}}_+$ (where $\underline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{0\}$) is a random variable with cdf $F(u) = \Pr(U \leq u)$ that is assumed continuous, independent of x, but dependent possibly on unknown parameters that are collected into a vector δ_u. Likewise, the error component $V = v \in \mathbb{R}$ has cdf $G(v) = \Pr(V \leq v)$ that is assumed to be continuous, independent of x, but dependent possibly on unknown parameters collected in vector δ_v. By Sklar's theorem, represent the joint cdf of U and V by

$$\begin{aligned} H(u, v) &= \Pr(U \leq u, V \leq v) \\ &= C_\theta(F(u), G(v)) \end{aligned}$$

where $C_\theta(\cdot, \cdot)$ is the bivariate copula of the joint distribution of U and V, that itself may depend on a vector of unknown parameters θ. It is assumed that θ has no elements in common with

$(\beta, \delta_u, \delta_v)$. Of course, the standard SFM arises when $C_\theta(x, y) = xy = \Pi$; that is, when U and V are assumed to be independent.

The likelihood function $L(\beta, \delta_u, \delta_v, \theta)$ is constructed from the distribution of the composite error $W = V - U$, where the random variable $W = w$ (where $-\infty < w < \infty$) is continuous. Denoting the pdf of W by $h_\theta(w)$, it can be derived by first considering the pdf of (U, V) :

$$\begin{aligned} h(u, v) &= \frac{\partial^2}{\partial u \partial v} H(u, v) \\ &= f(u) g(v) c_\theta(F(u), G(v)) \end{aligned} \tag{6}$$

where $f(u) = \partial F(u)/\partial u$ and $g(v) = \partial G(v)/\partial v$ denote, respectively, the pdf of U and the pdf of V, and c_θ is the copula density of C_θ. Note that the association between U and V is captured entirely by the density weighting function $c_\theta(F(u), G(v))$. Of course, if U and V are independent then $c_\theta(\cdot, \cdot) = 1$ implying that $h(u, v) = f(u)g(v)$, in which case the analyses presented in texts such as Coelli et al [3] and Kumbhakar and Knox Lovell [17] follow. Transforming $(U, V) \to (U, W)$ yields, as the pdf of (U, W) :

$$h(u, w) = f(u) g(u + w) c_\theta(F(u), G(u + w)). \tag{7}$$

Thus, the pdf of W is given by

$$h_\theta(w) = \int_{\underline{\boldsymbol{R}}_+} h(u, w) du \tag{8}$$

$$\quad = E_U \left(g(U + w) c_\theta(F(U), G(U + w)) \right) \tag{9}$$

where E_U denotes expectation with respect to the distribution of U. To illustrate the construction of the likelihood, consider here the simplest case: when there are a cross-section of n firms. Assuming independence across firms finds the likelihood given by

$$L(\beta, \delta_u, \delta_v, \theta) = \prod_{i=1}^{n} h_\theta(\log y_i - x_i' \beta)$$

where y_i is the output of firm i, and x_i its associated regressor vector.

It will be rare that a convenient closed form solution will exist to (8), although an exception is provided in Example 1 that follows, so optimisation of the likelihood will typically require numerical integration procedures. Should this prove onerous, an alternate approach is to use simulation estimation (cf. Greene [13]) as it is quite straightforward to cast $h_\theta(\cdot)$ in the form of an (unconditional) expectation with respect to U, see (9).

Of interest in the SFM is the technical efficiency TE_θ, that, following Battese and Coelli [1], can be measured by the following conditional expectation:

$$TE_\theta = E(\exp(-U) | W = w)$$

$$\quad = \frac{1}{h_\theta(w)} \int_{\underline{\boldsymbol{R}}_+} \exp(-u) h(u, w) du \tag{10}$$

$$\quad = \frac{E_U \left(\exp(-U) g(U + w) c_\theta(F(U), G(U + w)) \right)}{E_U \left(g(U + w) c_\theta(F(U), G(U + w)) \right)} \tag{11}$$

where (9) has been used to obtain (11). Once again, (11) can be used as the basis of simulation estimation of technical efficiency if the Laplace transform seen in (10) cannot be solved numerically, or (preferably) algebraically.

4 Examples

4.1 Example 1: Algebraic

In this example, closed form expressions for the pdf $h_\theta(w)$ and the technical efficiency TE_θ are derived for the Logistic-Exponential SFM. Underpinning this particular SFM is the fact that simple relationships exist between the pdf and cdf of the Logistic, as they also do for the pdf and cdf of the Exponential, thereby rendering much of the analysis of the Logistic-Exponential SFM attractively simple. Moreover, as the pdf of the Logistic has a shape similar to that of the pdf of the Normal, replacing the Normal by the Logistic may not cause too great discrepancies in the theory. For the copula associating the error components, the Fairlie-Gumbel-Morgenstern (FGM hereafter) family of copulas is selected. Assume:

- **Assumption #1:** The technical inefficiency error component U is Exponential with parameter $\sigma_u \geq 0$, with pdf

$$\begin{aligned} f(u) &= \frac{1}{\sigma_u}\exp\left(-\frac{u}{\sigma_u}\right) \\ &= \frac{1}{\sigma_u}(1-F(u)) \end{aligned} \tag{12}$$

  where

$$F(u) = 1 - \exp\left(-\frac{u}{\sigma_u}\right) \tag{13}$$

  is the cdf of U. It is well-known that $E(U) = \sigma_u$ and $Var(U) = \sigma_u^2$.

- **Assumption #2:** The noise error component V is Logistic with mean zero and scale parameter $\sigma_v > 0$, with pdf

$$\begin{aligned} g(v) &= \frac{1}{\sigma_v}\exp\left(-\frac{v}{\sigma_v}\right)\left(1+\exp\left(-\frac{v}{\sigma_v}\right)\right)^{-2} \\ &= \frac{1}{\sigma_v}G(v)(1-G(v)) \end{aligned} \tag{14}$$

  where

$$G(v) = \left(1+\exp\left(-\frac{v}{\sigma_v}\right)\right)^{-1} \tag{15}$$

  is the cdf of V. The scale parameter σ_v relates to variance according to

$$Var(V) = \frac{\pi^2}{3}\sigma_v^2.$$

- **Assumption #3:** The copula of the joint distribution of U and V is a member of the FGM family of copulas:

$$C_\theta(x,y) = xy(1+\theta(1-x)(1-y))$$

  where the association parameter θ is such that $-1 \leq \theta \leq 1$; clearly $C_0(x,y) = \Pi$. The copula density of the FGM family is given by

$$c_\theta(x,y) = 1 + \theta(1-2x)(1-2y). \tag{16}$$

Under these assumptions it is easy to show that

$$Cov(U,V) = \frac{1}{2}\theta\sigma_u\sigma_v$$

and thus the composite error $W = V - U$ of the Logistic-Exponential SFM is such that

$$E(W) = -\sigma_u \tag{17}$$

and

$$Var(W) = \left(\sigma_u^2 + \frac{\pi^2}{3}\sigma_v^2\right) - \theta\sigma_u\sigma_v. \tag{18}$$

When U and V are independent (i.e. when $\theta = 0$), $Var(W)$ is given by the term in braces. Moreover, as would be expected from the functional form $W = V - U$, if U and V are positively associated (i.e. when $\theta > 0$) then $Var(W)$ is reduced, while if they are negatively associated then $Var(W)$ is inflated.

To derive $h_\theta(w)$, begin with the pdf of (U, W) that is obtained by substituting (12)-(16) into (7)

$$\begin{aligned} h(u,w) &= \frac{1}{\sigma_u\sigma_v}\left.(1-x)y(1-y)(1+\theta(1-2x)(1-2y))\right|_{x\to F(u),y\to G(u+w)} \\ &= \frac{1}{\sigma_u\sigma_v}\left(y - xy - y^2 + xy^2 + \theta y - 3\theta xy + 2\theta x^2 y - 3\theta y^2\right. \\ &\qquad \left.\left. + 9\theta xy^2 - 6\theta x^2y^2 + 2\theta y^3 - 6\theta xy^3 + 4\theta x^2y^3\right)\right|_{x\to F(u),y\to G(u+w)} \end{aligned} \tag{19}$$

which is clearly a polynomial in $F(u)G(u+w)$. Consequently, to derive $h_\theta(w)$ as per (8) is straightforward. Before giving the general result, consider the leading case corresponding to independence (i.e. when $\theta = 0$) with pdf

$$\begin{aligned} h_\Pi(w) &= \frac{1}{\sigma_u\sigma_v}\int_0^\infty \left.\left(y - xy - y^2 + xy^2\right)\right|_{x\to F(u),y\to G(u+w)} du \\ &= \frac{1}{\sigma_u\sigma_v}\int_0^\infty \exp\left(-\frac{u}{\sigma_u}\right)\exp\left(-\frac{u+w}{\sigma_v}\right)\left(1+\exp\left(-\frac{u+w}{\sigma_v}\right)\right)^{-2} du \\ &= \frac{z}{\sigma_u}\int_0^1 t^{\sigma_v/\sigma_u}(1+zt)^{-2}dt \\ &= \frac{z}{\sigma_u+\sigma_v}\,{}_2F_1\left(2, 1+\frac{\sigma_v}{\sigma_u}; 2+\frac{\sigma_v}{\sigma_u}; -z\right) \end{aligned} \tag{20}$$

where

$$z = \exp\left(-\frac{w}{\sigma_v}\right).$$

Just as it occurs in $h_\Pi(w)$, the general solution for $h_\theta(w)$ that follows involves the Gaussian hypergeometric function ${}_2F_1(\cdot)$. In general form it is given by

$$ {}_2F_1\,(a,b;c;s) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-st)^{-a}dt \tag{21}$$

$$ = \sum_{i=0}^{\infty}\frac{(a)_i(b)_i}{(c)_i}\frac{s^i}{i!} \tag{22}$$

provided $\mathrm{Re}(c) > \mathrm{Re}(b) > 0$, and where the Pochhammer symbol $(a)_i = a(a+1)...(a+i-1)$. The further conditions required for its existence illustrate how the function should be computed. The integral form (21) is due to Euler, being single-valued and analytic provided $|\arg(1-s)| < \pi$.

The series form (22) converges under the stricter condition $|s| < 1$. Both forms are needed when computing the particular ${}_2F_1(\cdot)$ that appears in (20): the series form (22) can be used when $w > 0$, and when $w \leq 0$, the integral form (21) must be used. Well written computer code should switch automatically between the two forms as required; as does, for example, the Mathematica® computer algebra system (see Wolfram [29]). For further details on the Gaussian hypergeometric function see, for example, Slater [27] and Luke [18].

For arbitrary θ, the pdf $h_\theta(w)$ is obtained by substituting (19) into (8) and integrating term by term:

$$\begin{aligned} h_\theta(w) &= z\left[\frac{1-\theta}{\sigma_u+\sigma_v}\,{}_2F_1\left(2, 1+\frac{\sigma_v}{\sigma_u}; 2+\frac{\sigma_v}{\sigma_u}; -z\right)\right. \\ &\quad +\frac{2\theta}{\sigma_u+2\sigma_v}\,{}_2F_1\left(2, 1+\frac{2\sigma_v}{\sigma_u}; 2+\frac{2\sigma_v}{\sigma_u}; -z\right) \\ &\quad +\frac{2\theta}{\sigma_u+\sigma_v}\,{}_2F_1\left(3, 1+\frac{\sigma_v}{\sigma_u}; 2+\frac{\sigma_v}{\sigma_u}; -z\right) \\ &\quad \left.-\frac{4\theta}{\sigma_u+2\sigma_v}\,{}_2F_1\left(3, 1+\frac{2\sigma_v}{\sigma_u}; 2+\frac{2\sigma_v}{\sigma_u}; -z\right)\right]. \end{aligned} \tag{23}$$

The mean and variance of W derived respectively in (17) and (18) can now be verified directly from (23). The pdf $h_\theta(w)$ is illustrated in Figure 1, where, fixing $\sigma_u = 2$ and $\sigma_v = 1$, the three pdf curves plotted correspond to $\theta = -1, 0, 1$.

(Figure 1 about here.)

As the values $\theta = -1, 1$ lie at the extremes of θ-space, then, for fixed σ_u and σ_v, all pdfs $h_\theta(w)$ for any chosen θ such that $-1 < \theta < 1$ will be contained within the indicated curves, as quite clearly is the pdf $h_\Pi(w)$. While the mean is the same across all three pdf (here equalling -2), the variances differ, equalling 9.29 when $\theta = -1$, 7.29 when $\theta = 0$, and 5.29 when $\theta = 1$. All three pdf display skew to the left, consistent with negative values for Pearson's skewness measure $\sqrt{\beta_1}$ (here equalling -0.78 when $\theta = -1$, -0.81 when $\theta = 0$, and -0.82 when $\theta = 1$). However, upon inspecting its formula:

$$\sqrt{\beta_1} = \frac{3\sqrt{3}\sigma_u^2\left(3\theta\sigma_v - 4\sigma_u\right)}{2\left(3\sigma_u^2 - 3\theta\sigma_u\sigma_v + \pi^2\sigma_v^2\right)^{3/2}}$$

it is apparent that some parameter configurations can force $\sqrt{\beta_1} > 0$, this occurring if $\theta > \frac{4}{3}\sigma_u/\sigma_v \leq 1$. It should be noted that if $\sqrt{\beta_1} > 0$, then this need not invalidate (12)-(16) as a proper SFM, for despite the sign of $\sqrt{\beta_1}$, $h_\theta(w)$ need not be skewed to the right. This can be contrasted against the opinion expressed by Kumbhakar and Knox Lovell [17, p. 73] that attributes positive skewness as indicative only of model misspecification.

Next, consider the technical efficiency TE_θ defined in (10). For the leading case it is given by

$$\begin{aligned} TE_\Pi &= \frac{z}{h_\Pi(w)}\int_0^\infty \exp\left(-u\left(1+\frac{1}{\sigma_u}+\frac{1}{\sigma_v}\right)\right)\left(1+\exp\left(-\frac{u+w}{\sigma_v}\right)\right)^{-2} du \\ &= \frac{\sigma_u+\sigma_v}{\sigma_u+\sigma_v+\sigma_u\sigma_v}\frac{{}_2F_1\left(2, 1+\sigma_v+\frac{\sigma_v}{\sigma_u}; 2+\sigma_v+\frac{\sigma_v}{\sigma_u}; -z\right)}{{}_2F_1\left(2, 1+\frac{\sigma_v}{\sigma_u}; 2+\frac{\sigma_v}{\sigma_u}; -z\right)}. \end{aligned}$$

It is appealing that for fixed σ_v and w the ratio of hypergeometric functions appearing in TE_Π decreases monotonically in σ_u, as too does the initial scaling term. For θ non-zero, the technical efficiency is obtained by substituting (19) into (10) and integrating term by term:

$$
\begin{aligned}
TE_\theta \;=\; & \frac{z}{h_\theta(w)}\left[\frac{1-\theta}{\sigma_u+\sigma_v+\sigma_u\sigma_v}\;{}_2F_1\left(2,1+\sigma_v+\frac{\sigma_v}{\sigma_u};2+\sigma_v+\frac{\sigma_v}{\sigma_u};-z\right)\right.\\
& +\frac{2\theta}{\sigma_u+2\sigma_v+\sigma_u\sigma_v}\;{}_2F_1\left(2,1+\sigma_v+\frac{2\sigma_v}{\sigma_u};2+\sigma_v+\frac{2\sigma_v}{\sigma_u};-z\right)\\
& +\frac{2\theta}{\sigma_u+\sigma_v+\sigma_u\sigma_v}\;{}_2F_1\left(3,1+\sigma_v+\frac{\sigma_v}{\sigma_u};2+\sigma_v+\frac{\sigma_v}{\sigma_u};-z\right)\\
& \left.-\frac{4\theta}{\sigma_u+2\sigma_v+\sigma_u\sigma_v}\;{}_2F_1\left(3,1+\sigma_v+\frac{2\sigma_v}{\sigma_u};2+\sigma_v+\frac{2\sigma_v}{\sigma_u};-z\right)\right].
\end{aligned}
$$

In Figure 2, TE_θ is plotted against σ_u, where $\sigma_v = 1$ and $w = -0.1$, the three curves correspond to $\theta = -1, 0, 1$.

(Figure 2 about here.)

Interestingly, while TE_θ in this graph is initially larger for negative θ than it is for positive θ, this ordering is reversed once σ_u becomes larger than approximately 3.6. However, such switching does not necessarily occur at other settings for σ_v, w and θ in this SFM. A further important feature that can be seen in Figure 2 is that TE_θ is monotonically decreasing in σ_u; this can be evidenced at other settings for σ_v, w and θ, and would appear to be a property of the correlated error component Logistic-Exponential SFM. Efficiency monotonicity with respect to σ_u is a desirable property of an SFM, and is one that it is not necessarily shared even by some standard SFM models. For example, the TE_Π measure of technical efficiency in the Normal-Truncated Normal SFM (e.g. see Kumbhakar and Knox Lovell [17, eq. (3.2.52)]) is not everywhere monotonic decreasing in the scale parameter of the inefficiency error component.

4.2 Example 2: Cross-Section Data

In this example, correlated error component cost functions are fitted to a cross-section of $n = 123$ firms sampled from the US electricity utility industry (these data are listed in Greene [10, Table 3]). The models to be fit are cost functions rather than production functions, and specified as:

$$
\log\left(\frac{Cost}{P_f}\right) = \beta_0 + \beta_1 \log Q + \beta_2 \log^2 Q + \beta_3 \log\left(\frac{P_l}{P_f}\right) + \beta_4 \log\left(\frac{P_k}{P_f}\right) + V + U
$$

where output Q is a function of labour, capital and fuel, with respective factor prices P_l, P_k and P_f (cf. Greene [10, eq. (46)]). For the margins, assume:

- **Assumption #1:** The cost efficiency error component U is Half-Normal with scale parameter $\sigma_u \geq 0$, its pdf is $f(u) = 2\phi(u/\sigma_u)/\sigma_u$ and its cdf $F(u) = 2\Phi(u/\sigma_u) - 1$, where $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ denotes the pdf of a $N(0,1)$ random variable and $\Phi(x)$ denotes the corresponding cdf.

- **Assumption #2:** The noise error component V is Normal with scale parameter $\sigma_v > 0$, its pdf is $g(v) = \phi(v/\sigma_v)/\sigma_v$ and its cdf $G(v) = \Phi(v/\sigma_v)$.

Although the components of $W = V + U$ are now added, the theory given in Section 3 goes through largely unscathed. The necessary mathematical changes require replacing "$u + w$" in (7) with "$w - u$", and "$U + w$" in (9) and (11) with "$w - U$".

(Table 1 about here.)

The bivariate copula families to be examined here are the AMH family, the Frank family, and the Plackett family; see Table 1 for details of their functional forms and associated densities, parameter spaces and Spearman's S_ρ coefficient. The other bivariate copula to be considered is the Product copula Π corresponding to independence between U and V, yielding the standard (unassociated) Normal-Half-Normal SFM. The Product copula is nested within each of the copula families: for the AMH, set $\theta = 0$; for the Frank, allow $\theta \to 0$; and for the Plackett, allow $\theta \to 1$. Obviously, the AMH, the Frank and the Plackett are non-nested with respect to θ. Symbol θ is used generically and is not comparable across copula families. For a measure of association that is comparable, Spearman's S_ρ can be used; it can be seen in Table 1 that S_ρ is a (typically non-linear) function of θ. For the AMH family, S_ρ is bounded to the interval $[-0.2711, 0.4784]$. As the Frank and the Plackett families are comprehensive, S_ρ is not subject to any restriction outside of its natural bound $[-1, 1]$.

For all models, ML estimation is performed using the BFGS algorithm, with the estimators asymptotic variance-covariance estimated by the final iterate of the approximation to the inverse Hessian generated at each step of the algorithm. As there is no analytic solution to (8) for the correlated error component models considered here, numerical integration is used in order to form the log-likelihood and the score vector for each of these models. Programming utilised Mathematica® (see Wolfram [29]) as it provides a seamless means of transition from the algebra of the model to its numerical implementation. The ML estimates are reported in Table 2.

(Table 2 about here.)

In the first column appears the ML estimates for the standard (unassociated) Normal-Half-Normal SFM. The remaining columns list the ML estimates for the correlated error component models. While fairly comparable in magnitude, the estimates of $(\beta_0, ..., \beta_4)$ exhibit some variation across the estimated models, with perhaps the most marked differences occurring across the estimates of β_0 and of β_4. Not surprisingly, most of the impact of the correlated error component models can be seen in the variance structure (parameters σ_u, σ_v and θ), with the estimates of σ_u and σ_v increasing in magnitude for the copula models. The estimates of θ are consistent with negative association between the error components for these data, although the estimated standard error on each is obviously fairly large. However, due to the non-linear relationships between θ and S_ρ for these copula families (see Table 1), the parameter transformation from θ to Spearman's S_ρ yields, in the case of the Plackett and the Frank, significant negative estimates of association between U and V for these data. For completeness, other re-parameterisations of σ_u and σ_v that have been proposed in the SFM literature are reported in the second half of Table 2.

As none of the copula models given in Table 2 are nested (ignoring for the moment the Product copula), model selection amongst these can be based on measures such as AIC and BIC, following the suggestion of Joe [14, Sec.10.3]. Moreover, because the margins are fixed across competing models (i.e. models compete only according to copula specification and so the number of parameters does not vary across models), then selection criteria that use information measures that penalise fit by the number of parameters used to attain that fit are equivalent simply to selection based on the largest of the maximised log-likelihoods. On this basis, preferred here is the SFM estimated using the Frank family of copulas. In addition, for these data preferred overall is the Frank SFM, with an approximate p-value of 2.1% for the usual asymptotic χ^2 test for the nested comparison between it and the standard (unassociated) Normal-Half-Normal SFM; observed is $-2(66.1383 - 68.7975) = 5.32$ on a one degree of freedom test.

(Table 3 about here.)

Turning to cost efficiency TE_θ, it is estimated for each firm by replacing w in (10) with the observed residual $\widehat{w}$, and integrating numerically (except for the Product copula for which a closed form solution exists, see Kumbhakar and Knox Lovell [17, eq. (4.2.14)]). Table 3 gives point estimates for the first and last five-ranked firms for the standard (unassociated) Normal-Half-Normal SFM (i.e. using the Product copula) and the preferred Frank Normal-Half-Normal

SFM. Evidently, there is a dramatic difference between the values of the point estimates; for example, the cost efficiency of "Carolina P. & L." under the Frank SFM is estimated at 82.75%, differing considerably from the corresponding estimate of 94.49% under the standard SFM. The estimate difference is further magnified for the least efficient firms. In terms of cost efficiency ranking, there is broad agreement between both models (the empirical Spearman rank correlation coefficient equals 0.9636), especially as to those firms that are least cost efficient. However, at the upper end there is some variation in identifying the most cost efficient firms. For example, "Carolina P. & L." is first-ranked under the Frank SFM, but under the standard SFM it is ranked eighth. While the firm ranked first under the standard SFM is "New Mex. Elec. Ser.", but under the preferred Frank SFM this firm is ranked very much lower at fifty-sixth.

Admittedly there is a great deal of variation associated with efficiency estimates, but the reduced estimates seen for the firms listed in Table 3 occurs uniformly across all firms in the sample. This is evidenced by Figure 3, that plots the kernel smoothed distributions of cost efficiency estimates (used in both cases is the Epanechnikov kernel with bandwidth set to 0.06).

(Figure 3 about here.)

For these data, the distributional assumption of association between the error components demonstrates sensitivity across competing models in the measurement of firm technical efficiency, and consequently in the efficiency ranking of firms; cf. Coelli et al. [3, p.187, footnote 4].

4.3 Example 3: Panel Data

In this example, a correlated error component production function is fitted to an unbalanced panel of $n = 10$ firms sampled from the US airline industry over $T = 15$ years. These annual data are listed in Greene [11, pp.683-685] and described in Greene [12, Sec. 6.2]; it is assumed that the firms match the first 10 airline companies named in [12, Table 3.1]. The model is

$$\log Y = \beta_0 + \beta_1 \log E + \beta_2 \log F + \beta_3 \log L + \beta_4 \log M + \beta_5 \log P + V - U$$

where output Y is a function of inputs: equipment E, fuel F, labour L, materials M and property P. In observational terms, assume that the composite error $W_{it} = V_{it} - U_{it}$ follows the random-effects formulation of Kumbhakar [16] with time-varying technical efficiency:

$$W_{it} = V_{it} - \beta_t U_i \tag{24}$$

(i.e. $U_{it} = \beta_t U_i$) where

$$\beta_t = \left(1 + \exp(\gamma_1 t + \gamma_2 t^2)\right)^{-1}$$

with γ_1 and γ_2 real-valued parameters to be estimated ($i = 1, ..., n$ and $t = 1, ..., T$). For the purposes of this example, specify the margins as follows:

- **Assumption #1:** The efficiency error component U is Truncated Normal, i.e. a Normal with mean μ and scale parameter $\sigma_u \geq 0$ that is left-truncated at zero, with pdf $f(u) = \phi((u-\mu)/\sigma_u)/\sigma_u \Phi(\mu/\sigma_u)$ and cdf $F(u) = 1 - \Phi((\mu - u)/\sigma_u)/\Phi(\mu/\sigma_u)$.

- **Assumption #2:** The noise error component V is Normal with scale parameter $\sigma_v > 0$, with pdf $g(v) = \phi(v/\sigma_v)/\sigma_v$ and cdf $G(v) = \Phi(v/\sigma_v)$.

Given the margins, constructing a model for intra-firm correlated error components that is common to all firms requires the specification of a suitable $T + 1$ dimensional copula, or $(T+1)$-copula for short. For the i^{th} firm, assume that the copula is to represent equi-association between U_i and V_{it} for $t = 1, ..., T$, in combination with serial independence amongst $V_{i1}, ..., V_{iT}$.

Such a dependence structure requires its 2-copula margin $C_\theta(x, y_t) = C_\theta(x, 1, ..., 1, y_t, 1, ..., 1)$ be functionally equivalent for all $t = 1, ..., T$, and its T-copula margin $C_\theta(1, y_1, ..., , y_T) = \prod_{t=1}^{T} y_t$. One candidate $(T + 1)$-copula is:

$$x \left(\prod_{t=1}^{T} y_t \right) \left(1 + \theta(1 - x)(T - \sum_{t=1}^{T} y_t) \right) \tag{25}$$

corresponding to a simplified version of the extended FGM family of copulas given by Johnson and Kotz [15]. In this case, the family is indexed by the association parameter θ, where $-1 \leq \theta \leq 1$, and can represent degrees of both positive and negative association; note that for each $t = 1, ..., T$, the 2-copula $C_\theta(x, y_t) = xy_t(1+\theta(1-x)(1-y_t))$ is the FGM copula (cf. Assumption 3 in Example 1). The leading case $\theta = 0$ yields the $(T+1)$-product copula underlying the standard SFM panel data models discussed in, for example, Kumbhakar and Knox Lovell [17, Sec. 3.3]. The copula density of (25) is given by

$$1 + \theta(1 - 2x)(T - 2 \sum_{t=1}^{T} y_t).$$

Given these specifications, the joint pdf of $(U_i, V_{i1}, ..., V_{iT})$ is given by the multivariate extension of (6):

$$h(u_i, v_{i1}, ..., v_{iT}) = f(u_i) \left(\prod_{t=1}^{T} g(v_{it}) \right) \left(1 + \theta(1 - 2F(u_i))(T - 2 \sum_{t=1}^{T} G(v_{it})) \right)$$

where the functions f, g, F, and G are specified in Assumptions 1 and 2 above.

A second candidate is the $(T + 1)$-Normal copula

$$\Phi_{T+1} \left(\Phi^{-1}(x), \Phi^{-1}(y_1), ..., \Phi^{-1}(y_T); \Omega \right) \tag{26}$$

where $\Phi_{T+1}(\cdot; \Omega)$ denotes the cdf of a $(T + 1)$-dimensional standard Normal variable with correlation matrix Ω. The dependence structure given above is specified through Ω. Set

$$\Omega = \begin{bmatrix} 1 & \theta\iota_T' \\ \theta\iota_T & I_T \end{bmatrix}$$

where ι_T denotes the $(T \times 1)$ vector of units, and I_T the $(T \times T)$ identity matrix. As $\det \Omega = 1 - \theta^2 T$, positive and negative association can be represented to the degree that $|\theta| \leq 1/\sqrt{T} = 0.2582$, as here $T = 15$. The copula density of (26) is given by

$$\frac{\phi_{T+1} \left(\Phi^{-1}(x), \Phi^{-1}(y_1), ..., \Phi^{-1}(y_T); \Omega \right)}{\phi \left(\Phi^{-1}(x) \right) \phi \left(\Phi^{-1}(y_1) \right) ... \phi \left(\Phi^{-1}(y_T) \right)}.$$

The particular structure for Ω (exploited by using in $\phi_{T+1}(\cdot; \Omega)$ the Cholesky decomposition of the adjoint of Ω) in tandem with Assumption 2 leads to considerable simplification in the expression for the joint pdf of $(U_i, V_{i1}, ..., V_{iT})$:

$$\begin{aligned} h(u_i, v_{i1}, ..., v_{iT}) &= f(u_i) \left(\prod_{t=1}^{T} g(v_{it}) \right) \frac{1}{\phi \left(\Phi^{-1}(F(u_i)) \right)} \\ &\quad \times \frac{1}{\sqrt{1 - \theta^2 T}} \phi \left(\frac{\Phi^{-1}(F(u_i)) - \theta \sum_{t=1}^{T} v_{it}/\sigma_v}{\sqrt{1 - \theta^2 T}} \right) \end{aligned}$$

corresponding to a special form of the Meta-Gaussian distribution, the latter a special case of the Meta-Elliptical class of distributions (e.g. see Fang et al [7]). It is obvious that the leading case $\theta = 0$ yields the $(T + 1)$-product copula underlying the standard SFM panel data models.

The procedure to derive the joint pdf of $(W_{i1},...,W_{iT})$ is identical to that undertaken in Section 3, apart from the change in dimensionality. Using the relations $v_{it} = w_{it} + \beta_t u_i$, cf. (24), finds the pdf given by

$$h_\theta(w_{i1},...,w_{iT}) = \int_{\underline{\boldsymbol{R}}_+} h(u_i, w_{i1},...,w_{iT}) du_i.$$

Finally, assuming independence across firms, the likelihood function can be written

$$L(\beta_0,...,\beta_5,\mu,\sigma_u,\sigma_v,\gamma_1,\gamma_2,\theta) = \prod_{i=1}^{n} h_\theta \left(\log y_{i1} - x'_{i1}\beta,...,\log y_{iT} - x'_{iT}\beta\right)$$

where y_{it} is the output of firm i in period t, with x_{it} the associated regressors; in this example there are 12 parameters to be estimated. If the panel is unbalanced, as it is in this example, then this can be handled by reducing the dimensionality of the copula. For example, if the observations on firm i in period t are missing, then set $y_t = 1$ in (25) and (26).

(Table 4 about here.)

In Table 4, ML estimates are given for the standard (unassociated), the Normal (Meta-Gaussian) and the FGM Normal-Truncated Normal SFMs. Note that for these data the optimised FGM Normal-Truncated Normal SFM occurs at the corner of the parameter space corresponding to $\theta = -1$, thus there is no standard error reported for fixed $\widehat{\theta} = -1$. Evidently, the (degenerate) FGM SFM fits these data better than either of the other two SFM, even achieving this with one fewer estimated parameter than the Normal SFM; consequently, it is the FGM SFM that is preferred here. This result finds for the presence of negative association between the error components for these panel data.

(Table 5 about here.)

In terms of efficiency TE_θ, estimates are given in Table 5 for the first time period in which all firms record data, this occurring when $t = 2$. While not uniformly greater, the estimates reported for the preferred FGM SFM appear on the whole to be slightly larger than those of the standard SFM. The efficiency estimates of "TWA", "American" and "Delta" show little difference for the FGM SFM, but these show a greater spread under the standard SFM. The dramatic shift in the distributions of fitted TE_θ seen previously in the cross-section example does not occur here, even though there is negative association between the error components. More interesting is the alteration in rankings when comparing the standard SFM to the FGM SFM. The preferred FGM SFM ranks "United" ahead of "PanAm", whereas the inferior standard SFM reverses this order. There is agreement in rankings between both fitted models as to the most inefficient firms. It should be noted that the rankings reported in Table 5 must hold throughout the sample period $t = 1,...,15$, as in neither model there are no firm-specific parameters specified.

5 Conclusion

This paper has revisited the standard stochastic frontier model with a view to allowing the data the opportunity to determine whether or not there exists any statistical association between the model's error components. Model construction proceeded according to the copula approach, the rationale being that this method facilitates the modelling of association between random variables, alongside the specification given for each margin. The copula approach permits the fitting of a suite of candidate models from which a preferred model can emerge based on using well-known information-theoretic criteria. The copula approach was applied to cross-section data and to unbalanced panel data. In the cross-sectional example, striking was the location shift in the distribution of the Battese-Coelli measure of technical efficiency, this due to the presence of statistically significant association between the error components. Finally, in both cross-section and panel data examples it was shown that, when present, error component association impacts to alter efficiency rankings.

References

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

[2] Carley, H., and Taylor, M. D. (2002). A new proof of Sklar's theorem. In Cuadras, C. M., Fortiana, J., and Rodriguez-Lallena, J. A. (eds.), *Distributions With Given Marginals*, Kluwer: Dordrecht, 29-34.

[3] Coelli, T., Prasada Rao, D. S., and Battese, G. E. (1998). *An Introduction to Efficiency and Productivity Analysis*, Kluwer: Boston.

[4] Dall'Aglio, G. (1991). Frechet classes: the beginnings. In Dall'Aglio, G., Kotz, S., and Salinetti, G. (eds.), *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, Kluwer: Dordrecht, 13-50.

[5] Dardanoni, V., and Lambert, P. (2001). Horizontal inequity comparisons. *Social Choice and Welfare*, 18, 799-816.

[6] Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In Dempster, M. (ed.), *Risk Management: Value at Risk and Beyond*, Cambridge University Press: Cambridge, 176-223.

[7] Fang, H.-B., Fang, K.-T., and Kotz, S. (2002). The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82, 1-16.

[8] Fisher, N. I. (1997). Copulas. In Kotz, S., Read, C. B., and Banks, D. L. (eds.), *Encyclopedia of Statistical Sciences*, *Update Vol. 1*, Wiley: New York, 159-163.

[9] Frees, E. W., and Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2, 1-25.

[10] Greene, W. H. (1990). A gamma-distributed stochastic frontier model. *Journal of Econometrics*, 46, 141-163.

[11] Greene, W. H. (1995). *LIMDEP, Version 7.0 User's Manual*. Econometric Software, Inc: New York.

[12] Greene, W. H. (1997). Frontier production functions. In Pesaran, M. H., and Schmidt, P. (eds.), *Handbook of Applied Econometrics, Volume II: Microeconomics*. Blackwell: Oxford.

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

[14] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall: London.

[15] Johnson, N. L., and Kotz, S. (1975). On some generalized Farlie-Gumbel-Morgenstern distributions. *Communications in Statistics*, 4, 415-427.

[16] Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, 46, 201-212.

[17] Kumbhakar, S. C., and Knox Lovell, C. A. (2000). *Stochastic Frontier Analysis*. Cambridge University Press: Cambridge.

[18] Luke, Y. L. (1969). *The Special Functions and their Approximations*. Academic Press: New York.

[19] Miller, D. J., and Liu, W.-H. (2002). On the recovery of joint distributions from limited information. *Journal of Econometrics*, 107, 259-274.

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

[21] Nelsen, R. B. (1999). *An Introduction to Copulas.* Springer-Verlag: New York.

[22] Schweizer, B. (1991). Thirty years of copulas. In Dall'Aglio, G., Kotz, S., and Salinetti, G. (eds.), *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, Kluwer: Dordrecht, 13-50.

[23] Schweizer, B., and Sklar, A. (1983). *Probabilistic Metric Spaces.* North-Holland: New York.

[24] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institute de Statistique de l'Universite de Paris*, 8, 229-231.

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

[26] Sklar, A. (1996). Random variables, distribution functions, and copulas - a personal look backward and forward. In Rüschendorf, L., Schweizer, B., and Taylor, M. D. (eds.), *Distributions With Fixed Marginals And Related Topics.* Lecture Notes - Monograph Series, Volume 28, Institute of Mathematical Statistics: CA, 1-14.

[27] Slater, L. J. (1966). *Generalized Hypergeometric Functions.* Cambridge University Press: Cambridge.

[28] Smith, M. D. (2003). Modelling sample selection using Archimedean copulas. *The Econometrics Journal*, 6, 99-123.

[29] Wolfram, S. (1999). *The Mathematica Book*, 4th edition. Wolfram Media/Cambridge University Press: Cambridge.

Table 1: Copula Families

AMH	
Copula C_θ	$\frac{xy}{1-\theta(1-x)(1-y)}$
Density c_θ	$[1+\theta\,(xy+x+y-2+\theta\,(1-x)\,(1-y))]\,[1-\theta\,(1-x)\,(1-y)]^{-3}$
Parameter θ	$-1 \le \theta < 1$
Spearman's S_ρ (a)	$\frac{12(1+\theta)}{\theta^2}\,\mathrm{dilog}(1-\theta) - \frac{24(1-\theta)}{\theta^2}\log(1-\theta) - \frac{3(\theta+12)}{\theta}$
Frank	
Copula C_θ	$-\theta^{-1}\log\left(1+\frac{(e^{-\theta x}-1)(e^{-\theta y}-1)}{e^{-\theta}-1}\right)$
Density c_θ	$\theta\left(1-e^{-\theta}\right)e^{-\theta(x+y)}\left[1-e^{-\theta}-(e^{-\theta x}-1)(e^{-\theta y}-1)\right]^{-2}$
Parameter θ	$-\infty < \theta < \infty$
Spearman's S_ρ (b)	$1-\frac{12}{\theta}(D_1(\theta)-D_2(\theta))$
Plackett	
Copula C_θ (c)	$\frac{1}{2(\theta-1)}\,(s-t)$
Density c_θ	$\theta\,(s-2xy(\theta-1))\,t^{-3}$
Parameter θ	$\theta > 0$
Spearman's S_ρ	$\frac{\theta+1}{\theta-1}-\frac{2\theta}{(\theta-1)^2}\log\theta$
Normal	
Copula C_θ	$\Phi_2(\Phi^{-1}(x),\Phi^{-1}(y);\theta)$
Density c_θ	$\frac{\phi_2(\Phi^{-1}(x),\Phi^{-1}(y);\theta)}{\phi(\Phi^{-1}(x))\;\phi(\Phi^{-1}(y))}$
Parameter θ	$-1 \le \theta \le 1$
Spearman's S_ρ	$\frac{6}{\pi}\arcsin\left(\frac{\theta}{2}\right)$

Notes: (a) $\mathrm{dilog}(z)=\int_1^z \log(t)\,(1-t)^{-1}\,dt$ is the dilogarithm function.

(b) the Debye function $D_k(z)=kz^{-k}\int_0^z t^k\left(e^t-1\right)^{-1}dt$, for k any positive integer.

(c) $s=1+(\theta-1)(x+y)$ and $t=\sqrt{s^2-4xy\theta(\theta-1)}$.

Table 2: Energy Cost Parameter Estimates [(a)]

	Product copula	AMH copula	Plackett copula	Frank copula
β_0	-7.4083 (0.3185)	-7.5331 (0.3111)	-8.0261 (0.3846)	-8.0031 (0.3819)
β_1	0.4078 (0.0403)	0.4178 (0.0379)	0.4508 (0.0324)	0.4515 (0.0358)
β_2	0.0306 (0.0027)	0.0300 (0.0025)	0.0281 (0.0023)	0.0280 (0.0025)
β_3	0.2445 (0.0632)	0.2526 (0.0601)	0.3085 (0.0700)	0.2942 (0.0645)
β_4	0.0587 (0.0606)	0.0521 (0.0566)	0.0264 (0.0558)	0.0333 (0.0558)
σ_u	0.1553 (0.0476)	0.2098 (0.0611)	0.2937 (0.0793)	0.3589 (0.1428)
σ_v	0.1072 (0.0229)	0.1031 (0.0206)	0.1450 (0.0475)	0.1923 (0.1183)
θ		0.9291 (0.9009)	0.1077 (0.0973)	-8.3623 (6.6863)
$\log L$	66.1383	66.7551	68.4595	68.7975
S_ρ	0	-0.2550 (0.2064)	-0.6386 (0.1874)	-0.8159 (0.2184)
λ [(b)]	1.4497 (0.7292)	2.0339 (0.8880)	2.0263 (0.5499)	1.8668 (0.5658)
σ^2 [(b)]	0.0356 (0.0108)	0.0546 (0.0235)	0.1073 (0.0560)	0.1658 (0.1451)
γ [(b)]	0.6776 (0.2198)	0.8053 (0.1369)	0.8041 (0.0855)	0.7770 (0.1050)

Notes: (a) Figures shown to 4 dp. Estimated standard errors in parentheses.

(b) $\lambda = \sigma_u/\sigma_v$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\gamma = \sigma_u^2/\left(\sigma_u^2 + \sigma_v^2\right)$.

Table 3: Selected Energy Cost Efficiency Estimates and Ranks [a],[b]

Rank	Product copula			Frank copula		
1.	New Mex. Elec. Ser.	0.9705	(56. 0.7953)	Carolina P. & L.	0.8275	(8. 0.9449)
2.	Montana Power	0.9648	(52. 0.7992)	Central La. Pwr.	0.8274	(12. 0.9435)
3.	Northeast Util.	0.9590	(41. 0.8111)	Miss. Power & L.	0.8274	(15. 0.9408)
4.	Bangor Hydro.	0.9510	(18. 0.8254)	Community P.S.	0.8273	(7. 0.9481)
5.	Central Kansas	0.9504	(19. 0.8250)	Kansas Gas & El.	0.8273	(14. 0.9413)
119.	Colms. & So. Ohio	0.7601	(119. 0.6072)	Colms. & So. Ohio	0.6072	(119. 0.7601)
120.	Cal. Pac. Util	0.7339	(121. 0.4805)	United Gas. I.	0.5233	(121. 0.7180)
121.	United Gas. I.	0.7180	(120. 0.5233)	Cal. Pac. Util	0.4805	(120. 0.7339)
122.	N'western P.S.	0.7032	(122. 0.4766)	N'western P.S.	0.4766	(122. 0.7032)
123.	Maine Pub. Ser.	0.6786	(123. 0.3990)	Maine Pub. Ser.	0.3990	(123. 0.6786)
Mean [c]	0.8877			0.7677		
Std dev [c]	0.0543			0.0745		
Min [c]	0.6786			0.3990		
Max [c]	0.9705			0.8275		

Notes: (a) Figures shown to 4 decimal places.

(b) In parentheses is firm rank and efficiency estimate under the alternate copula.

(c) Statistics are for all $n = 123$ firms.

Table 4: Airline Production Parameter Estimates [a]

	Product copula	Normal copula	FGM copula
β_0	0.0973 (0.0283)	0.0856 (0.0291)	0.0537 (0.0177)
β_1	0.4827 (0.0917)	0.4835 (0.0837)	0.4360 (0.0708)
β_2	0.4336 (0.0554)	0.4402 (0.0562)	0.4969 (0.0494)
β_3	0.1974 (0.0467)	0.2041 (0.0512)	0.2223 (0.0458)
β_4	-0.1287 (0.0575)	-0.1501 (0.0652)	-0.1778 (0.0497)
β_5	0.0441 (0.0287)	0.0511 (0.0352)	0.0407 (0.0269)
μ	0.8396 (0.1162)	0.8635 (0.1248)	0.8040 (0.0861)
σ_u	0.1793 (0.0545)	0.1766 (0.0663)	0.1620 (0.0563)
σ_v	0.0662 (0.0047)	0.0657 (0.0045)	0.0663 (0.0042)
γ_1	0.0358 (0.0690)	0.0732 (0.0823)	0.0534 (0.0784)
γ_2	0.0279 (0.0092)	0.0262 (0.0119)	0.0305 (0.0110)
θ		-0.0719 (0.0845)	-1
$\log L$	152.2589	153.3257	154.3716
S_ρ [b]	0	-0.0687 (0.0808)	-0.3333
λ [c]	2.7085 (0.8638)	2.6884 (1.0497)	2.4440 (0.8838)
σ^2 [c]	0.0365 (0.0195)	0.0355 (0.0233)	0.0307 (0.0182)
γ [c]	0.8800 (0.0673)	0.8785 (0.0834)	0.8566 (0.0888)

Notes: (a) Figures shown to 4 dp. Estimated standard errors in parentheses.

(b) S_ρ concordance between U_i and V_{it}.

(c) $\lambda = \sigma_u/\sigma_v$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\gamma = \sigma_u^2/\left(\sigma_u^2 + \sigma_v^2\right)$.

Table 5: Year 2 Airline Production Efficiency Estimates, and Ranks [a]

Rank	Product copula		FGM copula	
1.	PanAm	0.7582	United	0.7733
2.	United	0.7399	PanAm	0.7683
3.	Continental	0.7268	Continental	0.7414
4.	Delta	0.7186	TWA	0.7118
5.	American	0.6883	American	0.7115
6.	TWA	0.6821	Delta	0.7079
7.	Eastern	0.6680	Eastern	0.6659
8.	National	0.6519	National	0.6582
9.	Braniff	0.6324	Braniff	0.6500
10.	Northwest	0.5846	Northwest	0.6006

Notes: (a) Figures shown to 4 decimal places.

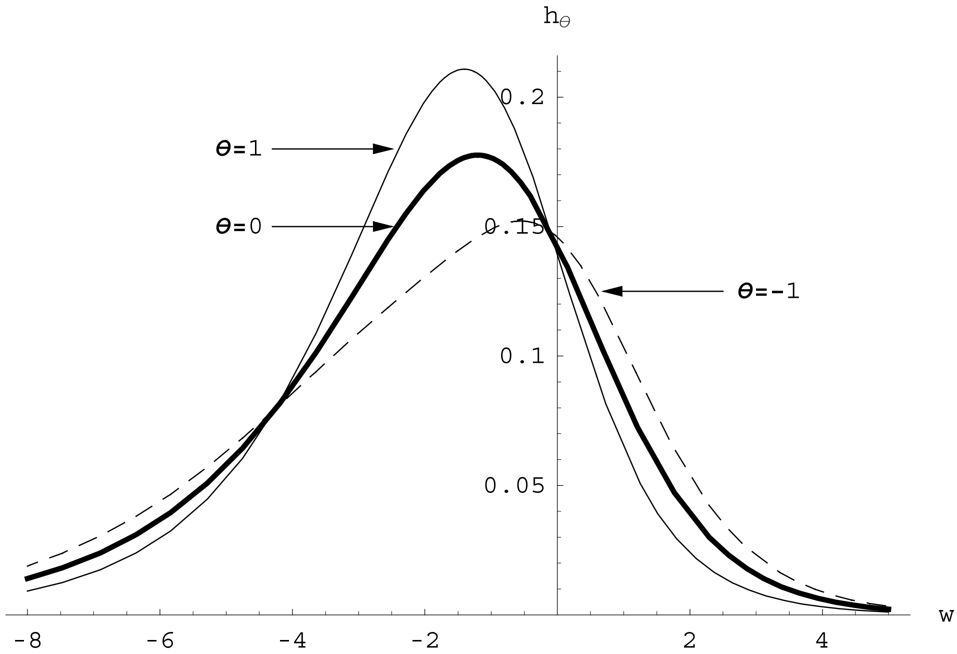

Figure 1: Probability density function $h_\theta(w)$ when $\theta = -1, 0, 1$, with $\sigma_u = 2$ and $\sigma_v = 1$

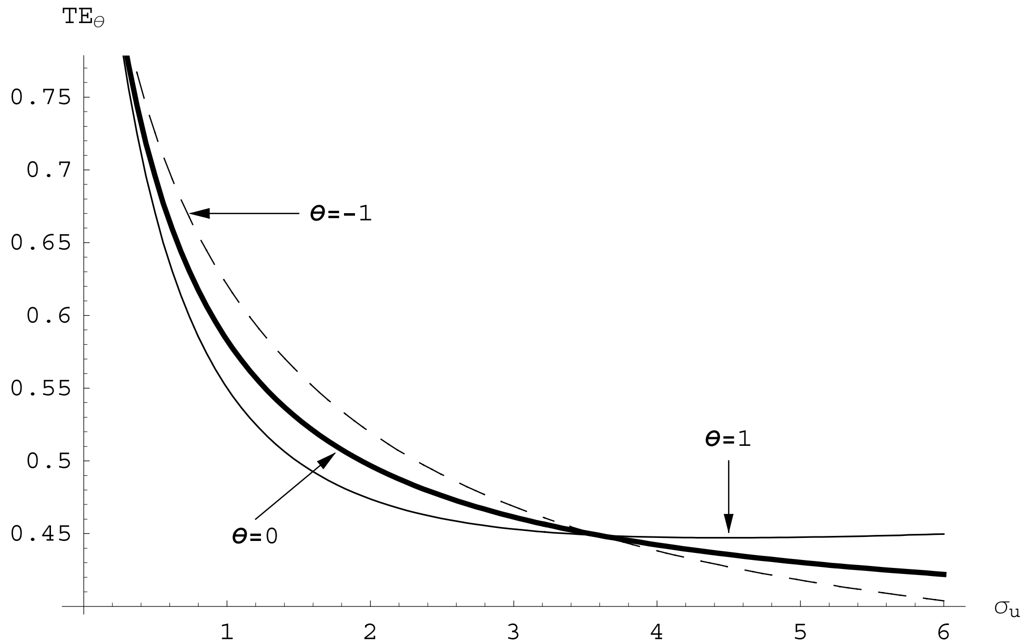

Figure 2: Technical efficiency TE$_\theta$ when $\theta = -1, 0, 1$, with $\sigma_v = 1$ and $w = -0.1$

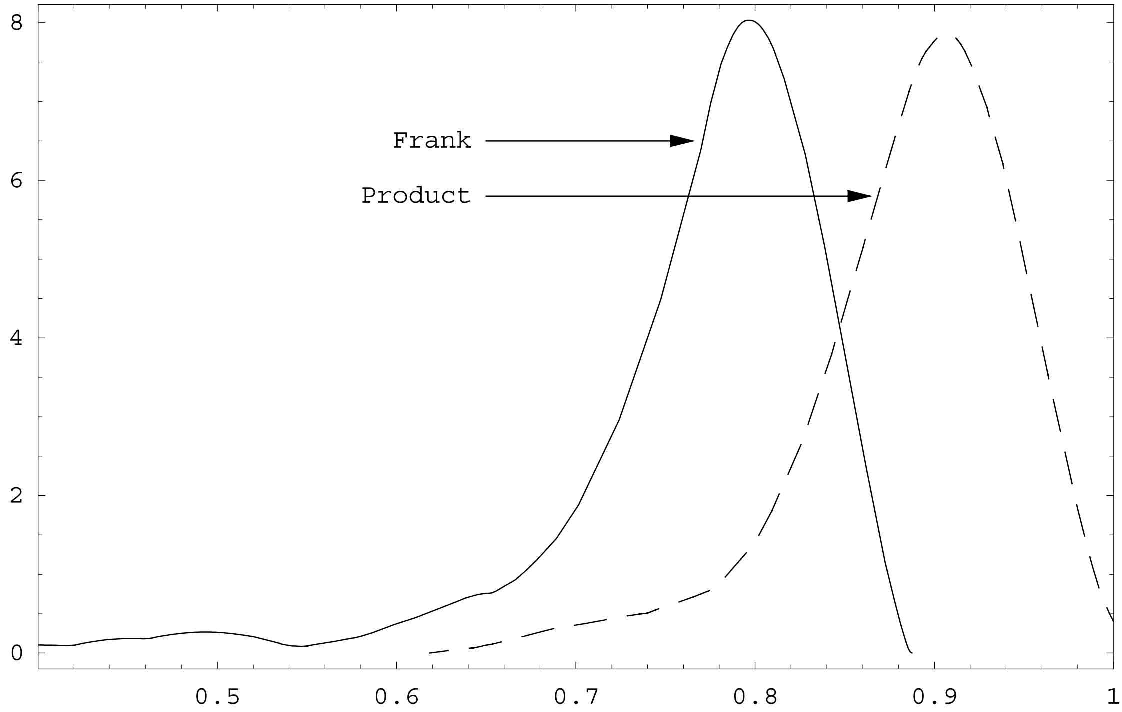

Figure 3: Kernel smoothed estimates of the distribution of Battese-Coelli cost efficiency