The 'wrong skewness' problem: a re-specification of stochastic frontiers

Graziella Bonanno [1] · **Domenico De Giovanni** [2] · **Filippo Domma** [2]

Abstract In this paper, we study the 'wrong skewness phenomenon' in stochastic frontiers (SF), which consists in the observed difference between the expected and estimated sign of the asymmetry of the composite error, and causes the 'wrong skewness problem', for which the estimated inefficiency in the whole industry is zero. We propose a more general and flexible specification of the SF model, introducing dependences between the two error components and asymmetry (positive or negative) of the random error. This re-specification allows us to decompose the third moment of the composite error into three components, namely: (i) the asymmetry of the inefficiency term; (ii) the asymmetry of the random error; and (iii) the structure of dependence between the error components. This decomposition suggests that the wrong skewness anomaly is an ill-posed problem, because we cannot establish ex ante the expected sign of the asymmetry of the composite error. We report a relevant special case that allows us to estimate the three components of the asymmetry of the composite error and, consequently, to interpret the estimated sign. We present two empirical applications. In the first dataset, where the classic SF has the wrong skewness, an estimation of our model rejects the dependence hypothesis, but accepts the asymmetry of the random error, thus justifying the sign of the skewness of the composite error. More importantly, we estimate a non-zero inefficiency, thus solving the wrong skewness problem. In the second dataset, where the classic SF does not yield any anomaly, an estimation of our model provides evidence for the presence of dependence. In such situations, we show that there is a remarkable difference in the efficiency distribution between the classic SF and our class of models.

1 Introduction

The basic formulation of a production stochastic frontier (SF) model[1] can be expressed as $y = f(x; \beta)e^{\mathcal{E}}$, where y is the firm production, x is a vector of inputs; β is vector of unknown parameters. The error term, $\mathcal{E} = V - U$, is assumed to be made of two statistically independent components, a positive random variable, denoted by U, and a symmetric random variable, denoted by V. U reflects the difference between the observed value of y and the frontier and can be interpreted as a measure of firms' inefficiency, while V captures random shocks, measurement errors and other statistical noise.

One major difficulty analysts often face when estimating an SF model is related to the choice of the distributions of

✉ Filippo Domma
filippo.domma@unical.it

[1] Department of Computer, Control and Management Engineering 'Antonio Ruberti', Sapienza University of Rome, Roma, Italy

[2] Department of Economics, Statistics and Finance, University of Calabria, Rende, Italy

---

[1] The original formulation of the SF model is based on the pioneering works of Aigner et al. (1977), Meeusen and van den Broek (1977) and Battese and Corra (1977) (see Kumbhakar and Lovell (2000) and Coelli et al. (2005), for a recent and comprehensive overview).

the random variables U and V. Different combinations have been proposed, including the normal—half-normal model (Aigner et al.1977; Battese and Corra1977), the normal—exponential model (Meeusen and van denBroek1977), the normal—truncated normal model (Stevenson 1980), and the normal—Gamma model (Greene 1990; Stevenson 1980). Perhaps the range of alternatives has been so far limited by computational challenges due to issues about the tractability of taking the convolution of the two error components. The choice of distributional specification is sometimes a matter of computational convenience.

The limited alternatives for the possible distributions also poses empirical challenges. For instance, several authors have addressed the problem related to the observed difference between the expected and the estimated sign of the asymmetry of the composite error. Specifically, for the standard SF model, the third central moment of $\mathcal{E}$ is

$$E\{[\mathcal{E} - E(\mathcal{E})]^3\} = -E\{[U - E(U)]^3\}, \tag{1}$$

thereby meaning, for example, that a positive skewness for the inefficiency term U implies an expected negative skewness for the composite error $\mathcal{E}$. However, in many applications, the residuals yield the wrong sign. This is called in the literature the wrong skewness phenomenon in SF models, initially pointed out by Green and Mayes (1991). In SF models, this phenomenon is important because it implies that the estimated variance of the inefficiency component is zero with positive probability. This, in turn, causes the inefficiency scores to be zero.

To overcome this problem, several authors have proposed the use, for the inefficiency component, of distribution functions with negative asymmetry. In particular, Carree (2002) uses the binomial probability function, Tsionas (2007) suggests the Weibull distribution whilst Qian and Sickles (2009),Almanidis and Sickles (2011) and Almanidis et al. (2014) consider a double truncated normal distribution.

More recent attempts to obtain the desired direction of the residual skewness are Feng et al. (2015), where the authors propose a finite sample adjustment to the existing estimators, and Hafner et al. (2016), where the authors propose a new approach to the problem by generalizing the distribution used for the inefficiency variable.

In this paper we argue that the wrong skewness problem has been only partially addressed because the relation described by Eq. (1), and the consequent discussion of the wrong skewness anomaly, is a direct consequence of all the assumptions underlying the specification of the basic formulation of an SF model. In fact, in a more general framework, where we relax the hypothesis of symmetry for V, of positive skewness for U, and of independence between U

and V, after simple but tedious algebra, the third central moment of the composite error turns out to be[2]

$$E\{[\mathcal{E} - E(\mathcal{E})]^3\} = -E\{[U - E(U)]^3\} + E\{[V - E(V)]^3\}$$
$$+ 3Cov(U^2, V) - 3Cov(U, V^2) - 6[E(U) - E(V)]Cov(U, V). \tag{2}$$

From Eq. (2), it is clear that the sign of the asymmetry of U and V and the dependence between U and V both affect the expected sign of the asymmetry of the composite error.

In order to take into account the different sources affecting the asymmetry of the composite error, in this paper we propose a very flexible specification of the SF model, introducing skewness in the random error V through a distribution whose shape can be asymmetric negative, positive, or symmetrical, depending on the value of one of its parameters, and the dependence between the two error components U and V. The dependence structure is modeled with a copula function that allows us to specify the joint distribution with different marginal probability density functions. Moreover, we use a copula function able to model positive or negative dependence, and the special case of independence, depending on the value of a dependence parameter.

In some special cases, the convolution between the two error components admits a semi-closed expression also in cases of statistical dependence between U and V. An example is provided in Smith (2008), who uses FGM copulas to relax the assumption of the independence of the two error terms. In a basic economic setting and with simple marginal distribution, Smith (2008) points out that the introduction of a statistical dependence between the two error terms may have a substantial impact on the estimated efficiency level. The author obtains an expression for the density of the composite error in terms of hypergeometric functions for a model with an exponential distribution for the inefficiency error and a standard logistic distribution for the random error. We propose a first generalization of Smith (2008) by using a Type I Generalized Logistic (GL) distribution for the random error. This distribution describes situations of symmetry or asymmetry (positive or negative), depending on the value of one of its parameters. This allows us to analyze the statistical properties of a model which has both statistical dependence and possible asymmetry in the random error component. While Kumbhakar and Lovell (2000) attribute some well-known limitations of the SF approach to incorrect specifications of the frontiers, we point out that some of the anomalies observed in the empirical literature may come from an incorrect specification of the shape of the density function of the two error components.

---

[2] The proof of this statement is available upon request.

A relevant point is the economic interpretation of our assumption about the dependence and asymmetry of the random error. As stressed by some authors, (Gómez-Déniz and Pérez-Rodriguez 2015; Pal and Sengupta 1999; Smith 2008), there are no statistical and economic reasons to assume the orthogonality of the errors. In particular, the independence of these error components in a productive industry is, in the opinion of Smith (2008), not obvious from either the economic or the statistical point of view. Pal and Sengupta (1999) argue that managerial efficiency may be affected by natural factors (representing statistical noise) in the agricultural sector. In addition, in a context in which current managerial decisions are influenced by past natural shocks (even in the short period), it easy to understand that the assumption of independence is too stringent. This holds particularly in agriculture, since a disturbance in one season will affect future decisions, thereby influencing the inefficiency component. Given that the same disturbance will affect the random error component, this could lead to the presence of a dependence between the random noise and the inefficiency. Also in this respect, Gómez-Déniz and Pérez-Rodriguez (2015) and Smith (2008) propose some examples of dependence in the agricultural sector. There are some good farmers who deal better with bad climatic conditions than do others. This will affect their efficiency when the weather is bad, meaning that a dependence among the error components cannot be excluded. A different interpretation is according to misspecification errors. In fact, there are some variables affecting the efficiency that cannot be included in the model because they are not well defined or not measurable (e.g., managerial ability, personal bias). Thus, they can flow into the statistical noise, given that it contains all things not included in the model, and this could lead to the presence of dependence (Pal and Sengupta 1999). In the manufacturing sector, unexpected events can occur in different phases of the productive process, influencing managerial decisions, e.g. damage to industrial machines, and defective products.

Regarding the asymmetry of the random error term, we depart from the huge literature which addresses this issue from a statistical point of view. In fact, some research deals with the normality of the residuals in different economic contexts (Azzalini 2005). From an economic perspective, unexpected shocks might affect the random error component, which does not necessarily follow a normal distribution. Normality seems to be a convenient assumption for mathematical computations. In several fields of research (economics, finance, sociology, engineering) error structures in regression models are not symmetric (Lewis and McDonald 2013; Lin et al.2013; Wu 2013). A more general example, in our opinion, would relate to a panel data context, thinking, for instance, macro-shocks affect the noise component of a production function. We expect that the skewness of this noise component could vary over time, following macro-cyclical patterns. For example, regarding our first empirical analysis of data from the NBER, we estimated production functions for different years and only in 19 years out of the 1958–2005 period were there cases of wrong skewness. We find that in one of these years (1979), the asymmetric random component contributes greatly to the decomposition of the third central moment. This dataset is a well known example of data evidencing wrong skewness (Bădin and Simar 2009; Hafner et al. 2016). The empirical results suggest that the dependence between the two error components is statistically rejected. Nonetheless, a strong positive asymmetry of the random error can be accepted. But, more importantly, our approach is able to estimate a non-zero inefficiency term, thus solving, at least in this dataset, the wrong skewness problem.

The second dataset has been chosen to test the class of models proposed in this paper when no wrong skewness is present in the data. In this empirical analysis, we find a strong non-linear dependence structure between the two sources of error. We show that this dependence structure has a significant impact on the observed Technical Efficiency (TE) scores.

Our model allows for statistical dependence through copulas in a straightforward manner. Thus, our paper is also related to the growing strand of the literature which uses copulas in stochastic frontiers. In particular, Amsler et al. (2014, 2016) introduce time dependence through copulas. Carta and Steel (2012) use copula functions to introduce a dependence between the outputs in a multi-output context. Lai and Huang (2013) propose a model taking into account the correlation among a set of individuals. Shi and Zhang (2011) use copulas for modeling the dependence in long-tail distributions. Tran and Tsionas (2015) model the dependence through a copula between endogenous regressors and the overall error term in a case in which external instruments are not available. We contribute to this literature by: (i) Giving a specification that generalizes Smith (2008); (ii) Indicating the exclusion of dependence as a possible cause of the wrong skewness phenomenon and the relative wrong skewnees problem; (iii) Suggesting that the presence of dependence between the two sources of error might considerably affect the estimation of TE scores.

This paper is organized as follows. In Section 2 we introduce the economic model and we list the steps required for the construction of the likelihood function and for the calculation of the TE. The new specification of the SF models is presented in Section 3, where a semi-closed expression for the probability density function of the model in terms of hypergeometric functions is derived. This allows us to discuss the statistical properties of the model in a rather transparent way. Section 4 reports the results of the two applications. In Section 5, we conclude. Then,

Appendix 1 presents the proof of our *Proposition 1*. Despite the semi-closed formula for the composite error function, the estimation of our examples requires a numerical discretization of the density. In this paper we use Gaussian quadratures, and the entire procedure is described in Appendix 2. Finally, Appendix 3 presents the analytical derivation of the TE and Appendix 4 presents the statistical properties of the copulas used in the empirical applications.

2 Stochastic frontiers and copula functions

The generic model of a production function for a sample of N firms is described as follows:

$$\overline{y} = \overline{x}\beta + V - U \tag{3}$$

where $\overline{y}$ is an $(N \times 1)$ vector of firms' log-outputs; $\overline{x}$ is an $(N \times K)$ matrix of log-inputs; β is a $(K \times 1)$ vector of unknown elasticities; V is an $(N \times 1)$ vector of random errors; U is an $(N \times 1)$ vector of random variables describing the inefficiencies associated to each firm (for a detailed discussion Kumbhakar and Lovell 2000).

To complete the description of the model, we need to specify the distributional properties of the random variables (U, V). The standard specification assumes independence between the random error and the inefficiency component, and a normal distribution for both random variables (although the inefficiency error must be truncated at zero to guarantee positiveness). We depart from this specification by considering a general joint probability density function $f_{U,V}(\cdot, \cdot, \Theta)$ for the couple (U, V), where Θ is the vector of parameters to be estimated, which includes β, the marginal and the dependence parameters. This density is defined on $\mathbb{R}^+ \times \mathbb{R}$, since inefficiency needs to be non-negative. The probability density function (pdf) of the composite error $\mathcal{E} := V - U$ is the convolution of two dependent random variables U and V, i.e.

$$f_{\mathcal{E}}(\epsilon) = \int_{\Re^+} f_{U,V}(u, \epsilon + u) du \tag{4}$$

where the joint probability density function, $f_{U,V}(u,v)$, is constructed using the properties of copula functions.

Copulas are widely appreciated tools used for the construction of joint distribution functions. To highlight the potential of this tool, it is sufficient to note that a copula function joins margins of any type (parametric, semi-parametric, and non-parametric distributions), not necessarily belonging to the same family, and captures various forms of dependence (linear, non-linear, tail dependence, etc.). A two-dimensional copula is a bivariate distribution function whose margins are uniform on (0,1). The importance of copulas stems from Sklar's theorem, which states how

copulas link joint distribution functions to their one-dimensional margins. Indeed, according to Sklar's theorem, any bivariate distribution $H(x,y)$ of the variables X and Y, with marginal distributions $F(x)$ and $G(y)$, can be written as $H(x,y) = C(F(x), G(y))$, where $C(.,.)$ is a copula function. Thus any copula, together with any marginal distribution, allows us to construct a joint distribution.

For the sake of parsimony, in this paper we do not include the rigorous construction of the copula function (details are Nelsen 1999). Rather, we describe the procedure we use to embed the copula into the stochastic frontier model described above (see also Smith (2008)), through 5 steps:

1. The choice of marginal distributions for the inefficiency error and the random error. We denote by $f_U(\cdot)$, $g_V(\cdot)$ their pdfs, and by $F_U(\cdot)$, $G_V(\cdot)$ their distribution functions.
2. The selection of the copula function $C_\theta(F_U(\cdot), G_V(\cdot))$. This usually involves additional dependence parameters, denoted here by θ.
3. The joint distribution function $f_{U,V}(u,v)$ has the following standard representation:

$$f_{U,V}(u, v) = f_U(u) g_V(v) c_\theta(F_U(u), G_V(v)), \tag{5}$$

   where $c_\theta(F_U(u), G_V(v)) = \frac{\partial^2 C_\theta(F_U(u), G_V(v))}{\partial F_U(u) \partial G_V(v)}$ is the density copula.
4. The pdf of the composite error $f_\epsilon(\cdot; \Theta)$ is the convolution of the joint density as in (4). Now, observing that $\epsilon_i = \overline{y}_i - \overline{x}_i\beta$, the likelihood function is given by

$$L = \prod_{i=1}^{N} f_{\mathcal{E}}(\overline{y}_i - \overline{x}_i\beta; \Theta) \tag{6}$$

   where $\overline{x}_i$ is the *ith* row of $\overline{x}$.
5. Finally, the technical efficiency (TE_Θ) is

$$TE_\Theta = E\left[e^{-U} | \mathcal{E} = \epsilon^*\right] = \frac{1}{f_{\mathcal{E}}(\epsilon^*; \Theta)} \int_{\Re^+} e^{-u} f_{U,V}(u, \epsilon + u; \Theta) du. \tag{7}$$

The complexity of the procedure described above depends on the choice of the marginal distribution functions $F_U(\cdot)$, $G_V(\cdot)$ and the copula function $C_\theta(\cdot, \cdot)$. It is equally obvious that the same choice influences the flexibility of the model. In the next section, we present a specification that represents a balanced trade-off between complexity and flexibility.

3 A new specification of SF models

In this section we propose a new specification for the composite error, which gives rise to a semi-closed expression for

Table 1 Marginal distribution functions and FGM copula

	Parameters	Probability density function	Distribution function
Exponential	$\delta_u > 0$	$\frac{1}{\delta_u} e^{-\frac{u}{\delta_u}}$	$1 - e^{-\frac{u}{\delta_u}}$
GL	$\alpha_v, \delta_v > 0$	$\frac{\alpha_v}{\delta_v} \frac{e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}}{\left(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{\alpha_v+1}}$	$\left(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{-\alpha_v}$
FGM	$\theta \in (-1,1)$	$1 + \theta(1-2F_u)(1-2G_v)$	$F_u G_v (1 + \theta(1-F_u)(1-G_v))$

the composite error of a production stochastic frontier. The density functions and the distribution functions of our specification are presented in Table 1. We model the inefficiency component of the composite error through an exponential random variable. This is a natural choice for the distribution of the inefficiency term, and has been used in many empirical applications (see Greene 1990; Smith 2008 to mention only a few). To model the random component of the composite error, we use a slight modification of the Type I generalized logistic distribution. To this end, we assume that the random variable V' is distributed as a Type I generalized logistic distribution with parameters $(\alpha_v, \delta_v, \lambda_v)$ and distribution function $G_{V'}(v') = \left(1 + e^{-\frac{v'-\lambda_v}{\delta_v}}\right)^{-\alpha_v}$, where α_v is a shape parameter, δ_v is a scale parameter, and λ_v is a location parameter (see, for example, Johnson et al. (1995)), and with expected value and variance given by $E[V'] = \lambda_v + \delta_v[\Psi(\alpha_v) - \Psi(1)]$ and $Var(V') = \delta_v[\Psi'(\alpha_v) + \Psi'(1)]$, respectively.

In order to interpret V' as a random error with zero mean, we consider the difference $V = V' - E(V')$. It is immediate to verify that the distribution function of V is $G_V(v) = \left(1 + e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{-\alpha_v}$, with $E(V) = 0$ and $Var(V) = Var(V')$.[3] Moreover, this choice makes our results directly comparable with those of Smith (2008), who uses a standard logistic distribution. Our results thus specialize to Smith (2008) when $\alpha_v = 1$. This particular choice for the distribution of the random error is motivated by the need for flexibility. Indeed, as shown in Domma (2004) and in Domma and Perri (2009), such s distribution is capable of generating either symmetric or positively/negatively skewed distributions. However, despite this flexibility, its density function is simple enough to be handled with relative ease. The final ingredient of our specification is the FGM copula. This assumption is motivated by the need to obtain a semi-closed form solution for the density of the random component. We would like to emphasize that in empirical applications, this copula is likely to produce poor estimates because of its well-known limitations. However, while at this stage the tractability is a priority, we will perform our empirical analyses with different specifications of the copula to check the robustness of our empirical results.

The following proposition presents a semi-explicit formulation of the pdf of the composite error in terms of a linear combination of hypergeometric functions,[4] the expected value, the variance and the third central moment of the composite error.

Proposition 1 *Assuming that $U \sim Exp(\delta_u)$, $V \sim GL(\alpha_v, \delta_v)$ and the dependence between U and V is modeled by an FGM copula. Let $k_1(\epsilon)$ be defined as $k_1(\epsilon) = exp\left\{-\frac{\epsilon+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}\right\}$*

1. *The density function of the composite error is*

$$f_{\mathcal{E}}(\epsilon; \Theta) = w_1(\epsilon) {}_2F_1\left(\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$+ w_2(\epsilon) {}_2F_1\left(\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$+ w_3(\epsilon) {}_2F_1\left(2\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$+ w_4(\epsilon) {}_2F_1\left(2\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right) \tag{8}$$

where the functions $w_1(.)$, $w_2(.)$, $w_3(.)$ and $w_4(.)$ are defined by

$$w_1(\epsilon) = (1-\theta)\frac{\alpha_v k_1(\epsilon)}{\delta_v + \delta_u} \quad w_2(\epsilon) = 2\theta\frac{\alpha_v k_1(\epsilon)}{2\delta_v + \delta_u}$$

$$w_3(\epsilon) = 2\theta\frac{\alpha_v k_1(\epsilon)}{\delta_v + \delta_u} \qquad w_4(\epsilon) = -4\theta\frac{\alpha_v k_1(\epsilon)}{2\delta_v + \delta_u}$$

2. *The expected value, the variance, and the third central moment of the composite error are*

$$E[\mathcal{E}] = -\delta_u, \tag{9}$$

---

[3] Throughout this paper, we will denote this distribution by GL(α_v, δ_v).

[4] The general form of a hypergeometric function is given by ${}_2F_1(a, b; c; s) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-st)^{-a}dt = \sum_{i=0}^{\infty} \frac{(a)_i(b)_i}{(c)_i}\frac{s^i}{i!}$ In the region $\{x : |s|<1\}$, it admits the following representation: ${}_2F_1(a, b; c; s) = \sum_{i=0}^{\infty}\frac{(a)_i(b)_i}{(c)_i}\frac{s^i}{i!}$ where $\Gamma(.)$ is the Gamma function and $(d)_i = d(d+1)\ldots(d+i-1)$ is the Pochhammer symbol, with $(d)_0 = 1$.

$$Var[\mathcal{E}] = \delta_u^2 + \delta_v^2[\Psi'(\alpha_v) + \Psi'(1)] - \frac{\theta\delta_u\delta_v}{2}[\Psi(2\alpha_v) - \Psi(\alpha_v)] \tag{10}$$

and

$$E\{[\mathcal{E} - E(\mathcal{E})]^3\} = -2\delta_u^3 + \delta_v^3[\Psi''(\alpha_v) - \Psi''(1)]$$

$$+ \frac{3}{4}\theta\delta_u\delta_v\{\delta_u[\Psi(2\alpha_v) - \Psi(\alpha_v)]$$

$$+ 2\delta_v[\Psi'(\alpha_v) + \Psi'(1)] - \delta_v[\Psi(2\alpha_v) \tag{11}$$

$$-\Psi(\alpha_v)]^2 - \delta_v[\Psi'(2\alpha_v) + \Psi'(1)]\}$$

where $\Psi(\cdot)$, $\Psi'(\cdot)$ and $\Psi''(\cdot)$ are, respectively, the Digamma, Trigamma and Tetragamma functions.

Proof See Appendix 1.

To appreciate the flexibility of our model, we point out that depending on the values of some parameters, we can specify the following four possible models:

- for $\theta = 0$ and $\alpha_v = 1$, we get the model of independence and symmetry, denoted by (I,S);
- for $\theta = 0$ and $\alpha_v \neq 1$, we have the model of independence and asymmetry, denoted by (I,A);
- for $\theta \neq 0$ and $\alpha_v = 1$, we obtain the model of dependence and symmetry, denoted by (D,S);
- for $\theta \neq 0$ and $\alpha_v \neq 1$, we have the model of dependence and asymmetry, denoted by (D,A)

In what follows, we will assess the impact of the asymmetry of the random error (via parameter α_v) and the effect of the dependence (via parameter θ) between U and V on the variance of the composite error. In particular, we compare four variances of the composite error, corresponding to the four models described above. First, we observe that for $\alpha_v = 1$, given that $\psi'(1) = \frac{\pi^2}{6}$ and $\psi(2)-\psi(1) = 1$, and by Eq. (10), we obtain the special case

$$Var_{\mathcal{E}}^{(D,S)} = \delta_u^2 + \frac{\pi^2}{3}\delta_v^2 - \frac{\theta\delta_u\delta_v}{2} \tag{12}$$

which overlaps Smith (2008) in the case of the symmetry of V and a dependence between U and V (this corresponds to a variance of $\mathcal{E}$ of the DS model). Moreover, to make the discussion simple, we highlight that the variance of the composite error in the cases of (a) independence and asymmetry and (b) independence and symmetry, are given, respectively, by $Var_{\mathcal{E}}^{(I,A)} = \delta_u^2 + \delta_v^2[\psi'(\alpha_v) + \psi'(1)]$ and $Var_{\mathcal{E}}^{(I,S)} = \delta_u^2 + \frac{\pi^2}{3}\delta_v^2$. Obviously, the variance of the composite error in the case of dependence and asymmetry is $Var_{\mathcal{E}}^{(D,A)} = Var(\mathcal{E})$ reported in Eq. (10).

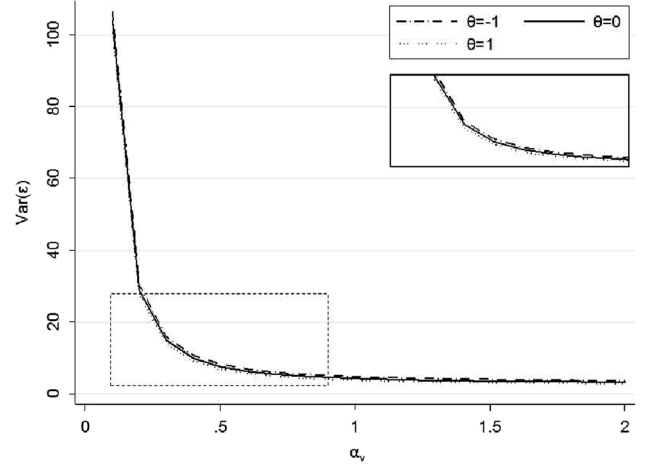**Fig. 1** Plot of $Var(\mathcal{E})$ for a production frontier with $\theta = -1$, $\theta = 0$ and $\theta = 1$ (α_v ranges between 0 and 2)

$$\cdots \alpha_v = 0.25; \quad — \alpha_v = 1; \quad --- \alpha_v = 3$$
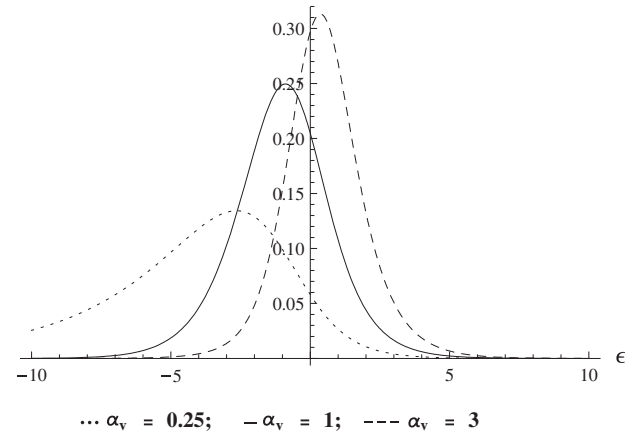
Fig. 2 Density function of $\mathcal{E}$ of a production frontier with $\delta_u = \delta_v = 1$, $\alpha_v = 1$ and $\theta = 1$ (α_v ranges between 0.25 and 3)

Figure 1 plots the variance of $\mathcal{E}$ as a function of α_v. The three lines corresponds to different dependence structures ($\theta = -1$, $\theta = 0$ or $\theta = 1$). In this figure, the effect of asymmetry on the variance of the composite error is particularly evident.

Next, we determine the effects of α_v on the distribution function of $\mathcal{E}$.[5] In fact, Fig. 2 shows how the asymmetry of the random error affects the distribution of the composite error. Imposing a maximal positive dependence between U and V ($\theta = 1$), we plot different pdfs for different values of α_v and observe that α_v affects not only the shape of the density, but also, and more importantly, the behavior of the distribution at the tails. The effect is more pronounced for negatively skewed distributions of the random error. This

---

[5] We have analyzed the impact of the dependence structure on the pdf of $\mathcal{E}$. Smith (2008) demonstrate this effect in the case of *symmetric-v*. We observe the same results under different conditions of skewness for V (negative or positive). For this reason, we do not show the plots here.

explains the impact on the variance observed above: negative skewness assigns much more probability mass to the extreme negative values of $\mathcal{E}$ than does positive skewness.

The empirical literature often finds estimated skewed density functions of the composite error which contrast with the theoretical predictions of the model (the 'wrong skewness' anomaly). In this regard, there is no general consensus on the interpretation of this misalignment between the assumptions and the observed facts. For instance, Kumbhakar and Lovell (2000) ascribe the misalignment to economically significant model misspecifications, while Simar and Wilson (2010) and Simar and Wilson (2011) claim that it may be due to an unfortunate sampling from a correctly specified population. Smith (2008) argues that the observed skewness may arise from the dependence between the random error and the inefficiency. Here, we contribute to the debate by suggesting one more possible explanation: it could be the interaction between the dependence (as argued by Smith) and the fundamental asymmetry of the distribution of the random error.

4 Empirical examples of production frontiers

In this section we present two empirical applications.[6] We use two different data samples, one in which a case of wrong skewness occurs, and one in which it does not occur.

For each dataset, we estimate the classic SF model and different variations within our class of models. For the reader's convenience, Table 2 presents the acronyms used to distinguish the models, together with their statistical properties. The acronyms have been chosen to mimic the statistical properties of the models analyzed, so that IS stands for Independence and Symmetry, IA for Independence and Asymmetry, and DA for Dependence and Asymmetry. IS is the most parsimonious, and is the specification closest to the classic SF. It cannot capture either asymmetries in the random component of the composite error or any dependence between the two error components.

In empirical applications, when selecting a copula function to capture some kind of dependence structure, one must take into account several factors, such as its tractability, the type of dependence (linear or non-linear), and the strength of the dependence. This last factor is usually measured with Kendall's τ_K (see Joe 1997; Nelsen 1999), which is defined as

$$\tau_K = 4 \int\int_{I^2} C(F_u, G_v) dC(F_u, G_v) - 1$$

---

for two arbitrary marginal U and V with distribution functions F_U and F_V coupled with the density copula $C(\cdot, \cdot)$. We test three different specifications of the copula function. The FGM copula (Nelsen 1999) (models DS and DA) has the advantage of producing a quasi-closed form density function for the composite error, and this allows us to obtain the decomposition in Proposition 1. However, in our applications this copula may suffer from some limitations. In particular, it can describe only situations where the dependence structure is limited, since it can only capture situations where $\tau_K \in [-2/9, 2/9]$. Stronger dependence structures need more sophisticated copulas. The bivariate Gaussian copula (included in DS_{Gauss} and DA_{Gauss} models) captures a linear dependence between two random variables. It is used mainly because it is easy to parametrize (Meyer 2013). It alleviates at least in part the limitations of the FGM, given that $\tau_K = \frac{2}{\pi} arcsin(\theta)$ where θ, in this case, is the correlation measure, but has the drawback of only being capable of capturing linear dependence. Lastly, the Frank copula (models DS_{Frank} and DA_{Frank}) is capable of capturing a non-linear dependence structure. It belongs to the family of Archimedean copulas and has a Kendall's $\tau_k = 1 - \frac{4}{\theta} + 4\frac{D_1(\theta)}{\theta}$, where $D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{x}{\exp^x - 1}dx$ (Huynh et al. 2014) and θ the association parameter. In Appendix 4, we present some of the mathematical characteristics of the Gaussian and Frank copula functions.

4.1 Wrong skewness in the data from the NBER database

We test our models in an SF production frontier using data from the NBER manufacturing productivity database (Bartelsman and Gray 1996). This archive is freely available online and contains annual information on US manufacturing industries from 1958 to the present.

In the underlying economic model, the variable 'value added' is our output, and total employment (*lemp*) and capital stock (*lcap*) are the input factors (all variables are in logarithms). The frontier assumes the Cobb-Douglas functional form.

We focus on the data for 1979 since we find the presence of a strong positive skewness in 1979, while negative skewness was expected from the traditional model (this was also shown in Hafner et al. (2016)). The wrong skewness phenomenon present in this data implies the wrong skewness problem in the sense that, according to the classical SF, the inefficiency hypothesis should be rejected.

4.1.1 Estimation

In this section we present the results in Tables 3 and 4, where significant coefficients are in bold (*t*-statistics are in parentheses).

Table 2 Summary of the statistical models

Name	Random error distribution	Inefficiency distribution	Dependence
Classic SF	Normal $\sim(0, \sigma_v^{\,2})$	Half-Normal $\sim(0, \sigma_u^2)$	No
IS	Symmetric GL $\sim(\alpha_v = 1, \delta_v)$	Exp $\sim(\delta_u)$	No
IA	GL $\sim(\alpha_v, \delta_v)$	Exp $\sim(\delta_u)$	No
DS	Symmetric GL $\sim(\alpha_v = 1, \delta_v)$	Exp $\sim(\delta_u)$	FGM copula
DA	GL $\sim(\alpha_v, \delta_v)$	Exp $\sim(\delta_u)$	FGM copula
DS$_{Gauss}$	Symmetric GL $\sim(\alpha_v = 1, \delta_v)$	Exp $\sim(\delta_u)$	Gaussian copula
DS$_{Frank}$	Symmetric GL $\sim(\alpha_v = 1, \delta_v)$	Exp $\sim(\delta_u)$	Frank copula
DA$_{Gauss}$	GL $\sim(\alpha_v, \delta_v)$	Exp $\sim(\delta_u)$	Gaussian copula
DA$_{Frank}$	GL $\sim(\alpha_v, \delta_v)$	Exp $\sim(\delta_u)$	Frank copula

What we observe first is a significant and strong positive asymmetry of the random error that, following the reasoning of this paper, should be the unique cause of the wrong skewness phenomenon (and the relative problem) of the data at hand.

Second, we observe that the DS and DA models (Table 3) provide non-significant association measures. In any case, testing the hypothesis $\theta = 0$ means testing whether there is independence between the inefficiency and the random component of the composite error. In this particular dataset, the data suggest that no dependence structure should be included.[7]

The Akaike Information Criterion (AIC) does not give a strong indication of which model should be preferred, since for the IA model the AIC is equal to -22.45, followed by IS (-21.15), DA (-20.52) and DS (-19.19).[8] However, even if each of our four models may be indifferent to the others (Burnham and Anderson 2004), the association measures are not significant and, comparing IA with IS, the former is preferred. Thus, the better fit goes in the direction of preferring models capturing asymmetry of the random error and not involving dependence structure.

As a robustness check, and driven by the high but not significant association parameters of the FGM copula, we have run a set of estimations in order to test whether other copulas can capture different forms of dependence. In more detail, we estimate both the DS and DA models, in which the dependence is captured through both the Gaussian and the Frank copulas. We present the results in Table 4. The main findings are: (i) If we look at the significance test on each estimated θ, the results suggest again the absence of dependence; (ii) Comparison of all the estimated models

suggests that the IA model remains the better fit, evidencing in a second way the absence of dependence in the NBER sample.

5 Asymmetry decomposition and technical efficiency

We now use the results of Proposition 1 to investigate the statistical aspects of the wrong skewness phenomenon and to decompose the asymmetry of the composite error.

Table 5 contains some descriptive statistics for the estimated parameters and the estimated composite error $\hat{\varepsilon}$. Each column represents one model, whose statistical characteristics are in Table 2. It is worth noting that the true direction of asymmetry is measured by factoring the sum of the deviations from the median (Zenga 1985).[9] One element of this decomposition is $E[\mathcal{E} - E(\mathcal{E})]^3$, which is the measure derived in Eq. (11). Table 5 presents the contributions of the single components to explain $E[\mathcal{E} - E(\mathcal{E})]^3$ and $E[\mathcal{E} - Me(\mathcal{E})]^3$.

We find a positive skewness of $\mathcal{E}$ in the IA and DA models, in which the v-component is strongly positive. The DS model also has a positive skewness due to the positive dependence-component. For the IS model, we can accept the symmetry of $\mathcal{E}$ (all the skewness measures are very close to 0). For the other specifications (IA, DS and DA) we have no wrong skewness anomaly. In fact, the signs of $E[\mathcal{E} - Me(\mathcal{E})]^3$ and $\sum[\hat{e} - Me(\hat{e})]^3$ are the same. We remark that: (i) IS assumes a priori that the v-component and the dependence-component are equal to 0, as in the classic SF; (ii) the dependence-component is positive for both models with a dependence structure, DA and DS, but it

---

[7] Note, however, that the fact that the association parameter θ in models DS and DA is extremely high suggests that the FGM copula does not account correctly for the dependence. In fact, Kendall's τ_K is probably well above the maximum level which the FGM copula can account for (0.22)

[8] Burnham and Anderson (2004) employ the measure $\Delta_i = AIC_i - AIC_{min}$ and consider that models having $\Delta_i \leqslant 2$ are supported with substantial evidence, those for which $4 \leqslant \Delta_i \leqslant 7$ have less support, and those for which $\Delta_{i>10}$ have no support.

[9] Departing from the demonstration of Zenga (1985) for descriptive measures, we obtain the following expression to account for the sign of the skewness:

$$E[\mathcal{E} - Me(\mathcal{E})]^3 = E[\mathcal{E} - E(\mathcal{E})]^3 + [E(\mathcal{E}) - Me(\mathcal{E})]^3 + 3[E(\mathcal{E}) - Me(\mathcal{E})]Var(\mathcal{E}). \tag{13}$$

Table 3 Estimations of production SF using US textile industry data (1979)

	Classic SF	IS	IA	DS	DA
β_0	**2.4507**	**2.4218**	**2.3969**	**2.5075**	**2.4117**
	(8.85)	(19.25)	(20.83)	(16.40)	(21.92)
β_1 (lemp)	**0.7920**	**0.7968**	**0.8323**	**0.7995**	**0.8306**
	(20.67)	(23.84)	(24.74)	(23.09)	(24.66)
β_2 (lcap)	**0.1913**	**0.1912**	**0.1815**	**0.1879**	**0.1821**
	(6.28)	(7.12)	(7.03)	(6.89)	(7.04)
δ_u		0.0002	0.0184	0.0750	0.0317
		(0.01)	(0.26)	(0.79)	(0.31)
α_v		–	**6.1212**	–	**6.0202**
			(33.46)		(50.20)
δ_v		**0.0994**	**0.1368**	**0.1035**	**0.1377**
		(8.63)	(7.61)	(3.48)	(3.24)
θ		–	–	0.99998	0.4063
				(0.35)	(0.07)
Obs	54	54	54	54	54
log-likelihood	13.58	15.58	17.22	15.59	17.26
AIC	−17.17	−21.15	−22.45	−19.19	−20.52

Source: our elaborations on data from the NBER productivity database. The dependent variable is the value added (in log). 5% significant coefficients are in bold

Table 4 Estimations of models capturing dependence through Gaussian and Frank copulas. NBER dataset

	DS_{Gauss}	DS_{Frank}	DA_{Gauss}	DA_{Frank}
β_0	**2.4804**	**2.4859**	**2.3978**	**2.4755**
	(20.33)	(16.74)	(19.20)	(10.66)
β_1 (lemp)	**0.7937**	**0.7944**	**0.8333**	**0.8148**
	(25.12)	(24.90)	(20.68)	(13.12)
β_2 (lcap)	**0.1873**	**0.1940**	**0.1809**	**0.1905**
	(7.37)	(7.29)	(6.85)	(3.22)
δ_u	**0.0132**	0.0788	0.0170	**0.0974**
	(3.00)	(0.90)	(1.34)	(2.58)
α_v	–	–	6.2964	7.6870
			(0.38)	(0.41)
δ_v	**0.0949**	**0.1169**	**0.1384**	**0.1879**
	(6.41)	(4.38)	(4.09)	(4.12)
θ	0.2677	5.0666	0.0588	11.4830
	(0.52)	(1.22)	(0.03)	(1.29)
Obs	54	54	54	54
log-likelihood	14.74	15.96	17.19	17.48
AIC	−17.48	−19.91	−20.38	−20.95

Source: our elaborations on data from the NBER productivity database. The dependent variable is the value added (in log). 5% significant coefficients are in bold

is very small, and also (iii) dependence is statistically rejected in this data sample.

Table 6 presents some descriptive statistics for the estimated TE for each model.[10]

The wrong skewness problem is evident in the classic SF and IS specifications. These models, indeed, present an industry where all manufacturers are perfectly efficient (the mean TE score is 1), and there is no variability (the variance of the efficiency score approaches zero). What we observe in our preferred model (IA) is that we can actually estimate a non-zero TE. In other words, the IA model has the ability to estimate the variability in the industry, thus solving, at least in this dataset, the wrong skewness problem.

5.1 Application to a sample of Italian manufacturing firms

We use data from AIDA ('*Analisi Informatizzata delle Aziende Italiane*'), which is a database containing financial and accounting information of Italian firms. This dataset does not evidence wrong skewness. Nonetheless, we would like to determine the impact of our specification on the distribution of the efficiency scores. It turns out, indeed, that incorporating dependence between the two sources of error significantly affects the estimated TE.

We use again a Cobb-Douglas production function where the dependent variable is the value added, representing the firms' output, while labor and capital are the traditional inputs. Moreover, we introduce ICT and R&D

---

[10] The derivation of the TE_Θ scores is presented in Appendix 3.

Table 5 Summary measures and skewness of the composite error (data from NBER)

	IS	IA	DS	DA
δ_u	0.0002	0.0184	0.0750	0.0317
α_v	1	6.1212	1	6.0202
δ_v	0.0994	0.1368	0.1035	0.1377
θ	0	0	0.99998	0.4063
$E(\epsilon)$	−0.0002	−0.0184	−0.0750	−0.0317
$V(\epsilon)$	0.0325	0.0344	0.0370	0.0340
$Me(\epsilon)$	−0.0009	−0.0432	−0.0744	−0.0562
$E[\epsilon-E(\epsilon)]^3$	−1.82E-11	0.006057	0.001575	0.006412
u-component	−1.82E-11	−0.000012	−0.000844	−0.000064
v-component	0	0.006069	0	0.006192
dependence-component	0	0	0.002419	0.000283
$E[\epsilon-Me(\epsilon)]^3$	−1.84E-06	0.0086	0.0015	0.0090
$\sum[\hat{\epsilon}-Me(\hat{\epsilon})]^3$	0.4513	0.5781	0.4498	0.4949

Source: our elaborations on data from the NBER productivity database

$E[\epsilon-Me(\epsilon)]^3 = E[\epsilon-E(\epsilon)]^3 + [E(\epsilon)-Me(\epsilon)]^3 + 3[E(\epsilon)-Me(\epsilon)]V(\epsilon)$

$E[\epsilon-E(\epsilon)]^3 = $ *u-component + v-component + dependence-component*

$\sum[\hat{\epsilon}-Me(\hat{\epsilon})]^3$ is calculated as shown in Zenga (1985) for descriptive measures

investments as additional inputs. All variables refer to the calendar year 2009 and are in logarithms.[11]

5.1.1 Estimation

The results from Table 7 highlight the robustness of the estimates across our models and the significance of all fitted parameters (*t*-statistics are in brackets). Moreover, in terms of the AIC, the classic SF models are very far from the other specifications. The distance is much more than 10 points (Burnham and Anderson 2004). In particular, the more general DA model has the better fit (AIC 1042.26), but the AIC for the DS model is very close (1042.46). Hence, we reject the asymmetry of the random error for all the specifications. Thus, the results indicate the presence of a positive dependence and the symmetry of *V* in this data sample.

The estimated values for θ give reasons to further investigate the dependence, but even more, the calculated values of τ_k convinced us to carry out a more in-depth analysis of the dependence (0.18 for the DS model and 0.22 for the DA model, considering that 0.22 is the maximum value for τ_k with the FGM copula). Table 8 presents the estimations of the DS and DA models capturing the dependence through the Gaussian and Frank copulas. If we look at the AIC, we can say that the better model is DS_{Frank}

(1034.47), also with respect to the estimations in Table 7. This means a strong positive dependence between *U* and *V*, in fact τ_k is equal to 0.48 (the maximum is 1).

5.1.2 Technical efficiency

In the previous subsection, we showed that from a purely statistical point of view, the dataset evidences a strong non-linear dependence between the inefficiency error and the random error. The questions are: Does this dependence affect the distribution of the efficiency scores? In which ways?

Table 9 displays the summary statistics of the efficiency scores for each of the 9 models under investigation (the classic SF as well as 8 models with different variants of our specification). We find significant differences in the observed distributions of the TE scores of the classic SF and our models. In particular, an inspection of Table 9 reveals that the left tails of the distributions are completely different. For instance, the classic SF gives a minimum efficiency level of 0.33, whereas the minimum level is 0.87 for DS_{Frank}, which is the best model in terms of AIC. We also observe a pronounced difference in terms of the expected values and the variability in the industry. To better highlight the difference in the distribution of the efficiency scores, we plot in Fig. 3 the kernel densities of the distributions of the efficiency scores observed in the classic SF and DS_{Frank}. The classic SF presents an industry where the efficiency is more variable among the firms than it is for DS_{Frank}. A number of firms in the industry are very inefficient: only a few firms are very efficient. On the other hand, DS_{Frank}

Table 6 Some descriptive statistics of Technical Efficiency (data from NBER)

	Classic SF	IS	IA	DS	DA	DS_{Gauss}	DS_{Frank}	DA_{Gauss}	DA_{Frank}
Mean	0.999546	0.9998	0.9819	0.9305	0.9693	0.9763	0.9725	0.9811	0.9781
Stand. Dev.	6.18E-07	2.26E-07	0.0022	0.0085	0.0034	0.0021	0.0046	0.0022	0.0103
Min	0.999545	0.9997	0.9738	0.9066	0.9543	0.9721	0.9599	0.9731	0.9483
Max	0.999548	0.9998	0.9840	0.9385	0.9709	0.9795	0.9765	0.9832	0.9895

Source: our elaborations on data from the NBER productivity database

Table 7 Estimations of production SF for the Italian manufacturing firms using data from AIDA (2009)

	Classic SF	IS	IA	DS	DA
β_0	**2.4104**	**2.3340**	**2.3232**	**2.4288**	**2.3737**
	(19.23)	(20.30)	(20.54)	(18.55)	(19.80)
β_1 (llabour)	**0.7224**	**0.7411**	**0.7420**	**0.7434**	**0.7447**
	(34.52)	(37.06)	(37.10)	(37.36)	(37.24)
β_2 (lcapital)	**0.0928**	**0.0944**	**0.0952**	**0.0946**	**0.0958**
	(7.24)	(7.74)	(7.87)	(7.82)	(7.98)
β_3 (lICT)	**0.1119**	**0.0991**	**0.0979**	**0.0980**	**0.0961**
	(7.55)	(6.98)	(6.94)	(6.95)	(6.82)
β_4 (lR&D)	**0.1035**	**0.0994**	**0.0997**	**0.0986**	**0.0985**
	(7.37)	(7.53)	(7.61)	(7.53)	(7.58)
δ_u		**0.1117**	**0.1047**	**0.1980**	**0.1376**
		(3.41)	(4.87)	(2.24)	(3.32)
α_v		−	**0.8819**	−	**0.8252**
			(8.76)		(7.32)
δ_v		**0.2234**	**0.2144**	**0.2299**	**0.2231**
		(26.48)	(19.35)	(10.69)	(14.58)
θ		−	−	**0.8032**	0.9990[a]
				(2.03)	(1.27)
Obs	937	937	937	937	937
log-likelihood	−536.27	−514.72	−514.15	−513.23	−512.13
AIC	1084.54	1043.44	1044.30	1042.46	1042.26

Source: our elaborations on data from the AIDA dataset. The dependent variable is the value added (in log)
[a] 20% significant level. 5% significant coefficients are in bold

presents a completely different picture of the industry. Here, a large number of firms are concentrated around the very high expected value, and only a few firms are moderately inefficient.

The observed differences in the distribution of the efficiency scores raise the question of whether the underlying economic model is correctly specified.

A possible explanation of the remarkable difference in the distributions of the TE is due to misspecification errors of the underlying economic model. In fact, some variables affecting the efficiency cannot be included in the model because they are either not well defined or not measurable. Thus, the bias in the economic model can flow into the statistical noise, given that this noise contains everything not included in the model, and this might introduce a dependence between the two different sources of error (see Pal and Sengupta 1999). In such situations, not properly taking into account the dependence structure might produce an estimated TE producing a different picture of the industry.

6 Conclusions

In this paper, we have shown that the so-called 'wrong skewness' anomaly in stochastic frontiers is a direct consequence of the basic hypotheses, which appear to be overly restrictive. In fact, relaxing the hypotheses of the symmetry of the random error and the independence of the components of the composite error, we obtain a re-specification of the stochastic frontiers model that is sufficiently flexible. This allows us to explain the difference between the

Table 8 Estimations of models capturing dependence through Gaussian and Frank copulas. AIDA dataset

	DS_{Gauss}	DS_{Frank}	DA_{Gauss}	DA_{Frank}
β_0	**2.2951**	**2.5573**	**2.3179**	**2.4125**
	(20.36)	(20.74)	(53.90)	(16.57)
β_1 (llabour)	**0.7377**	**0.7531**	**0.7362**	**0.7532**
	(36.89)	(38.42)	(54.94)	(37.85)
β_2 (lcapital)	**0.0953**	**0.0965**	**0.0967**	**0.0973**
	(7.81)	(8.32)	(8.95)	(8.32)
β_3 (lICT)	**0.0997**	**0.0917**	**0.0982**	**0.0912**
	(7.07)	(6.74)	(5.31)	(6.61)
β_4 (lR&D)	**0.1025**	**0.0966**	**0.1014**	**0.0953**
	(7.77)	(7.67)	(5.79)	(7.45)
δ_u	**0.1031**	**0.2887**	**0.1039**	**0.1361**
	(5.00)	(4.07)	(3.55)	(1.58)
α_v	–	–	**1.0092**	**0.8188**
			(8.46)	(6.60)
δ_v	**0.2208**	**0.2867**	**0.2203**	**0.2585**
	(25.98)	(8.12)	(25.25)	(15.86)
θ	−0.1621	**5.4569**	−0.1135	**4.7315**
	(−0.71)	(2.37)	(−1.07)	(3.04)
Obs	937	937	937	937
log-likelihood	−514.62	−509.23	−514.59	−509.47
AIC	1045.24	1034.47	1047.19	1036.93

Source: our elaborations on data from the AIDA database. The dependent variable is the value added (in log). 5% significant coefficients are in bold

Table 9 Some descriptive statistics of technical efficiency (data from AIDA)

	Classic SF	IS	IA	DS	DA	DS_{Gauss}	DS_{Frank}	DA_{Gauss}	DA_{Frank}
Mean	0.7638	0.8996	0.9054	0.8348	0.8791	0.9389	0.9180	0.9295	0.9494
Stand. Dev.	0.0794	0.0255	0.0221	0.0364	0.0060	0.0122	0.0092	0.0153	0.0080
Min	0.3371	0.8179	0.8448	0.6661	0.8688	0.9087	0.8677	0.8890	0.9273
Max	0.9137	0.9306	0.9342	0.8655	0.9124	0.9553	0.9255	0.9494	0.9580

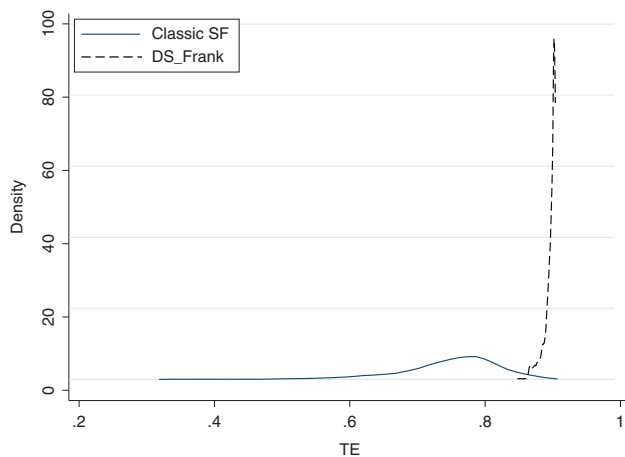*Source*: our elaborations on data from AIDA database

Fig. 3 Kernel density of efficiency scores for model classic SF and DS_{Frank}

expected and the estimated signs of the asymmetry of the composite error that is found in various applications of the classic stochastic frontier model.

We have decomposed the third moment of the composite error into three components, namely: (i) the asymmetry of the inefficiency term; (ii) the asymmetry of the random error; and (iii) the dependence structure of the error components. This enables us to reinterpret the unusual asymmetry in the composite error by measuring the contribution of each component in the model. This has been shown in one of the two empirical examples, i.e. the data from the NBER archive, for which a case of wrong skewness has been reported when using the classic SF specification.

When wrong skewness occurs, estimations with classic SF correspond to OLS estimations, and the inefficiency scores are zero. This misleads one to the conclusion that there is no inefficiency. Our specification allows

overcoming this difficulty, as witnessed in both empirical applications, where our estimates of the output elasticities with respect to the inputs are more robust than those involved in the standard SF specification and the estimated efficiency scores are lower than unity.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interests.

Appendix 1

1.1 Proof of proposition 1

In order to prove *Proposition 1* easily, we report some preliminary results in the following *Lemma*.

Lemma 1

1. $U \sim Exp(\delta_u)$ then

- *r*-th moment is $E(U^r) = \delta_u^r \Gamma(r+1)$. *Consequently, we have:* $E(U) = \delta_u$, $E(U^2) = 2\delta_u^2$ *and* $E(U^3) = 6\delta_u^3$.
- *Denoted with* $F(u) = 1 - e^{-\frac{u}{\delta_u}}$ *the distribution function of the random variable U, after algebra, we obtain* $E[U^r F(U)] = E(U^r)\left(1 - \frac{1}{2^{r+1}}\right)$

2. *If* $V \sim GL(\alpha_v, \delta_v)$, *with pdf* $g_V(v) = \frac{\alpha_v}{\delta_v} e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}$

$\left(1 + e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{-\alpha_v - 1}$ *then*

- $E(V) = 0$;
- $E(V^2) = Var(V) = \delta_v^2[\Psi'(\alpha_v) + \Psi'(1)]$;
- $E(V^3) = \delta_v^3[\Psi''(\alpha_v) - \Psi''(1)]$.
- *Denoted with* $G_V(v) = \left(1 + e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{-\alpha_v}$ *the distribution function of the random variable V, it is easy to prove that* $E[V^k G(V)] = (1/2)E\{[\delta_v(\Psi(2\alpha_v) - \Psi(\alpha_v)) + V]^k | 2\alpha_v, \delta_v\}$, *where* $E[.|2\alpha_v, \delta_v]$ *is the expectation with respect to the GL with parameters* $2\alpha_v$ *and* δ_v. *In particular, for* $k = 1$ *and* $k = 2$ *we have* $E[VG(V)] = \frac{\delta_v}{2}[\Psi(2\alpha_v) - \Psi(\alpha_v)]$ *and* $E[V^2 G(V)] = \frac{\delta_v^2}{2}\{[\Psi(2\alpha_v) - \Psi(\alpha_v)]^2 + [\Psi'(2\alpha_v) - \Psi'(1)]\}$, *respectively.*

3. *if* $(U,V) \sim f_{U,V}(u,v) = f_U(u)g_V(v)[1 + \theta(1-2F_U(u))(1-2G_V(v))]$ *then*

$$E(U^r V^k) = (1+\theta)E(U^r)E(V^k)$$

$$-2\theta\{E(U^r)E[V^k G(V)] + E[U^r F(U)]E(V^k) -$$

$$2E[U^r F(U)]E[V^k G(V)]\} = E(U^r)E(V^k) +$$

$$\theta\left(\frac{1}{2^r} - 1\right)E(U^r)\{E(V^k) - E(V^k G_V(V))\}$$

Now, we can prove the *Proposition* 1.

1. The *pdf* of composite error is $f_{\mathcal{E}}(\epsilon) = \int_{\Re^+} f_{U,V}(u, \epsilon+u)du$ where $f_{U,V}(u, \epsilon+u) = f_U(u)_V(\epsilon+u)c(F_V(u), G_V(\epsilon+u))$. Given that $c(\cdot, \cdot)$ is a density copula of a FGM copula, we have

$$f_{U,V}(u, \epsilon+u) = (1+\theta)f_U(u)g_V(\epsilon+u) - 2\theta f_U(u)$$

$$g_V(\epsilon+u)G_V(\epsilon+u) - 2\theta f_U(u)g_V(\epsilon+u)F_U(u)$$

$$+4\theta f_U(u)g_V(\epsilon+u)F_U(u)G_V(\epsilon+u) \tag{14}$$

Using (14), we have $f_{\mathcal{E}}(\epsilon) = (1+\theta)I_1 - 2\theta\{I_2 + I_3 - 2I\}$, where $I = \int_{\Re^+} f_U(u)g_V(\epsilon+u)F_U(u)G_V(\epsilon+u)du$, and I_i, for $i = 1,2,3$ are special cases of I.

Now, in order to calculate the integral I, we observe that

$$f_U(u)g_V(\epsilon+u)F_U(u)G_V(\epsilon+u)$$

$$= \frac{\alpha_v k_1(\epsilon)}{\delta_u \delta_v} e^{-\frac{u}{\delta_u} - \frac{u}{\delta_v}}\left(1 - e^{-\frac{u}{\delta_u}}\right)\left(1 + k_1(\epsilon)e^{-\frac{u}{\delta_v}}\right)^{-2\alpha_v - 1} \tag{15}$$

where $k_1(\epsilon) = e^{-\frac{\epsilon+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}$. After algebra, we can write

$$I = \frac{\alpha_v k_1(\epsilon)^{-2\alpha_v}}{\delta_u \delta_v}\left\{\int_{\Re^+}(e^{-u})^{\frac{1}{\delta_u}+\frac{1}{\delta_v}}[1 + k_1(\epsilon)(e^{-u})^{\frac{1}{\delta_v}}]^{-2\alpha_v-1}du -\right.$$

$$\left. - \int_{\Re^+}(e^{-u})^{\frac{2}{\delta_u}+\frac{1}{\delta_v}}[1 + k_1(\epsilon)(e^{-u})^{\frac{1}{\delta_v}}]^{-2\alpha_v-1}du\right\}$$

If before we put $y = e^{-u}$ and then $t = y^{\frac{1}{\delta_v}}$, after algebra, we obtain

$$I = \frac{\alpha_v k_1(\epsilon)}{\delta_u}\left\{\int_0^1 t^{\frac{\delta_v}{\delta_u}}(1 + k_1(\epsilon)t)^{-2\alpha_v-1}dt - \int_0^1 t^{2\frac{\delta_v}{\delta_u}}(1 + k_1(\epsilon)t)^{-2\alpha_v-1}dt\right\}$$

Bearing in mind that for hypergeometric function is true the following

$$\frac{\Gamma(c-b)\Gamma(b)}{\Gamma(c)} {}_2F_1(a, b; c; s) = \int\limits_0^1 t^{b-1}(1-t)^{c-b-1}(1-st)^{-a}dt$$

We obtain

$$I = \frac{\alpha_v k_1(\epsilon)}{\delta_u}\left\{\frac{1}{\frac{\delta_v}{\delta_u}+1} {}_2F_1\left(2\alpha_v+1, \frac{\delta_v}{\delta_u}+1; \frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\right)\right.$$

$$\left.-\frac{1}{2\frac{\delta_v}{\delta_u}+1} {}_2F_1\left(2\alpha_v+1, 2\frac{\delta_v}{\delta_u}+1; 2\frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\right)\right\}$$

2. By *Lemma 1*, we can to verify that

- $E(\epsilon) = -E(U) = -\delta_u$
- $Var(\epsilon) = Var(U) + Var(V) - 2cov(U,V) = \delta_u^2 + \delta_v^2$ $[\Psi'(\alpha_v) + \Psi'(1)] - 2cov(U,V)$, where $cov(U,V) = \frac{\theta}{2}E(U)E[VG_V(V)] = \frac{\theta}{4}\delta_u\delta_v[\Psi(2\alpha_v) - \Psi(\alpha_v)]$.
- Moreover, recalling that for a generic random variable, Z, we have $E[Z-E(Z)]^3 = E(Z^3) - 3E(Z^2)E(Z) + 2[E(Z)]^3$, after simple algebra, $E[U-E(U)]^3 = 2\delta_u^3$ and $E[V - E(V)]^3 = \delta_v^3[\Psi''(\alpha_v) - \Psi''(1)]$. Moreover, by Lemma, we have:

$$cov(U^2, V) = E[U^2V] = \frac{3}{4}\theta\delta_u^2\delta_v[\Psi(2\alpha_v) - \Psi(\alpha_v)]$$

and

$$cov(U, V^2) = -\frac{\theta}{4}\delta_u\delta_v^2\{2[\Psi'(\alpha_v) + \Psi'(1)] - [\Psi(2\alpha_v)$$

$$-\Psi(\alpha_v)]^2 - [\Psi'(2\alpha_v) + \Psi'(1)]\}$$

by (2), after algebra, we obtain $E[\epsilon - E(\epsilon)]^3$ as in Eq. (11).

Appendix 2

2.1 The numerical procedure

The estimation of models like those described in Section 2 requires the ability to compute the density of the composite error. Closed-form expressions for this quantity are available only in some few special cases, such as the noteworthy case addressed by Smith (2008). While in the previous section we provided one more example of a closed-form expression, this section is intended to describe the scheme we use to approximate the likelihood (6) starting from a general joint density $f_{U,V}$. Our goal is to provide a numerical tool capable of managing different joint distributions for the couple (U,V), thus widening the set of alternatives one can use when defining an SF model.

Our approach is fairly simple. We approximate the convolution of U with V by means of a numerical quadrature. To be more precise, put $\mathcal{E} = V - U$. Its density function, $f_{\mathcal{E}}(\cdot; \Theta)$, is obtained by the convolution of U with V:

$$f_{\mathcal{E}}(x; \Theta) = \int\limits_0^\infty f_{U,V}(u, x+u; \Theta)du \qquad (16)$$

An explicit evaluation of the integral in (16) is in general infeasible, which has kept a potential range of possible joint densities almost unexplored. However, an approximation of (16) by Gauss-Laguerre quadrature has proved to be easy and effective, and is presented below.

Let us first rewrite (16) as

$$f_{\mathcal{E}}(x; \Theta) = \int\limits_0^\infty e^{-u}g_x(u)du, \qquad (17)$$

with $g_x(u) = e^u f_{U,V}(u, x+u)$. Fix an integer m, which we will refer to as the order of quadrature. For $h = 1\ldots,m$, let: (i) t_h be the h-th root of the Laguerre polynomial of order m, $L_m(u)$, and (ii) ω_h be defined by the following system of linear equations[12],[13]

$$\int\limits_0^\infty s^k e^{-s}ds = \sum_{h=1}^n \omega_h t_h^k \quad k = 1, \ldots, 2m-1. \qquad (18)$$

Then, we can write

$$f_{\mathcal{E}}(x; \Theta) \approx \sum_{h=1}^m \omega_h g_x(t_h). \qquad (19)$$

Inasmuch as the function $g_x(\cdot)$ is Riemann integrable on the interval $[0,\infty)$, standard results in numerical analysis ensure the goodness of the approximation.

We can thus approximate the integral appearing in (16) (and its gradient with respect to Θ) by a finite sum, and insert the approximating density function and its gradient into a quasi Newton-like iteration (however, from experience with the normal—half-normal model with the FGM copula, a few initial iterations with the algorithm of Berndt et al. (1974) is highly recommended). As for the order of quadrature, practice with the normal—half-normal case with the FGM copula shows that $m = 12$ is sufficient to obtain *safe* approximations. For values of m around 12, computations of the Laguerre nodes and weights require a fraction of a second, and this is needed only once.

---

[12] The system is over-determined, but possesses a unique solution $\omega_1, \ldots, \omega_n$.

[13] These are basic concepts in numerical analysis. For more details about orthogonal polynomials and Gaussian quadrature, any textbook in this topic may be consulted. A standard reference for economists is Judd (1998).

Appendix 3

3.1 Calculation of TE scores

Given the *Proposition* 1, its Proof in Appendix 1 and Eq. (7) in Section 2, we derive the formula to calculate the Technical Efficiency scores TE_Θ for our models.

We can write

$$TE_\Theta = E\left[e^{-U}|\mathcal{E} = \epsilon^*\right] = \frac{1}{f_\mathcal{E}(\epsilon^*; \Theta)} \int_{\Re^+} e^{-u} f_{U,V}(u, \epsilon + u; \Theta) du$$

(20)

where $f_{U,V}(u, \epsilon + u)$ is derived in Eq. (14).

After algebra, we obtain:

$$TE_\Theta = E[e^{-U}|\mathcal{E} = \epsilon]$$

$$= \frac{\overline{\omega}_1(\epsilon)\overline{H}_1(\epsilon) + \theta\left[\overline{\omega}_1(\epsilon)\overline{H}_1(\epsilon) - 2\overline{\omega}_2(\epsilon)\overline{H}_2(\epsilon) - 2\overline{\omega}_3(\epsilon)\overline{H}_3(\epsilon) + 4\overline{\omega}_4(\epsilon)\overline{H}_4(\epsilon)\right]}{\omega_1(\epsilon)H_1(\epsilon) - \theta\left[\omega_1(\epsilon)H_1(\epsilon) - 2\omega_2(\epsilon)H_2(\epsilon) - 2\omega_3(\epsilon)H_3(\epsilon) + 4\omega_4(\epsilon)H_4(\epsilon)\right]}$$

(21)

where the $H-functions$ represent hypergeometric functions. In particular, we have:

$$\overline{H}_1 = {}_2F_1\left(\alpha_v + 1, \frac{\delta_v}{\delta_u} + \delta_v + 1; \frac{\delta_v}{\delta_u} + \delta_v + 2; -k_1(\epsilon)\right)$$

$$H_1 = {}_2F_1\left(\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$\overline{H}_2$$

$$= \frac{1}{\left[\delta_v\left(\frac{1}{\delta_u} + 1\right)\right]} {}_2F_1\left(\alpha_v + 1, \delta_v\left(\frac{1}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{1}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right)$$

$$- \frac{1}{\left[\delta_v\left(\frac{2}{\delta_u} + 1\right)\right]} {}_2F_1\left(\alpha_v + 1, \delta_v\left(\frac{2}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{2}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right)$$

$$H_2 = {}_2F_1\left(\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$\overline{H}_3 = {}_2F_1\left(2\alpha_v + 1, \frac{\delta_v}{\delta_u} + \delta_v + 1; \frac{\delta_v}{\delta_u} + \delta_v + 2; -k_1(\epsilon)\right)$$

$$H_3 = {}_2F_1\left(2\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

$$\overline{H}_4$$

$$= \frac{1}{\left[\delta_v\left(\frac{1}{\delta_u} + 1\right)\right]} {}_2F_1\left(2\alpha_v + 1, \delta_v\left(\frac{1}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{1}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right)$$

$$- \frac{1}{\left[\delta_v\left(\frac{2}{\delta_u} + 1\right)\right]} {}_2F_1\left(2\alpha_v + 1, \delta_v\left(\frac{2}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{2}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right)$$

$$H_4 = 2F_1\left(2\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)$$

with $k_1(\epsilon) = e^{-\frac{\epsilon + \delta_v[\Psi(\alpha_v) - \Psi(1)]}{\delta_v}}$ and the $\omega-functions$ are respectively defined as:

$$\overline{\omega}_1(\epsilon) = \overline{\omega}_3(\epsilon) = \frac{\alpha_v k_1(\epsilon)}{\delta_u\left[\delta_v\left(\frac{1}{\delta_u} + 1\right)\right]} \quad \overline{\omega}_2(\epsilon) = \overline{\omega}_4(\epsilon) = \frac{\alpha_v k_1(\epsilon)}{\delta_u}$$

$$\omega_1(\epsilon) = \omega_3(\epsilon) = \frac{\alpha_v k_1(\epsilon)}{\delta_v + \delta_v} \qquad \omega_2(\epsilon) = \omega_4(\epsilon) = \frac{\alpha_v k_1(\epsilon)}{2\delta_v + \delta_v}$$

Appendix 4

4.1 Gaussian and Frank copula functions

	Gaussian	Frank
Parameter	$\theta \in (-1, 1)$	$\theta \in (-\infty, +\infty) \backslash \{0\}$
Density	$\frac{1}{\sqrt{1 - \theta^2}} exp\left(\frac{2\theta\Phi^{-1}[F(u)]\Phi^{-1}[G(v)] - \theta^2\left(\Phi^{-1}[F(u)]^2 + \Phi^{-1}[G(v)]^2\right)}{2(1 - \theta^2)}\right)$	$\frac{\theta\left(1 - e^{-\theta}\right)e^{-\theta(F(u) + G(v))}}{[(1 - e^{-\theta}) - (1 - e^{-\theta F(u)})(1 - e^{-\theta G(v)})]^2}$
Distribution	$\Phi(\Phi^{-1}[F(u)], \Phi^{-1}[G(v)]; \theta)$	$-\theta^{-1}\ln\left[1 + \frac{\left(e^{-\theta F(u)} - 1\right)\left(e^{-\theta G(v)} - 1\right)}{(e^{-\theta} - 1)}\right]$

Legend: Φ is the Standard Normal distribution and Φ^{-1} is the inverse function.

References

Aigner D, Lovell C, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econom 6:21–37

Almanidis P, Sickles RC (2011) The skewness issue in stochastic frontiers models: fact or fiction? In: Van Keilegom I, Wilson PW (eds) Exploring research frontiers in contemporary statistics and econometrics. Springer, Berlin, pp 201–227

Almanidis P, Qian J, Sickles RC (2014) Stochastic frontier models with bounded inefficiency. In: Sickles RC, Horrace WC (eds) Festschrift in honor of Peter Schmidt: econometric methods and applications. Springer, New York, NY, pp 47–81

Amsler C, Prokhorov A, Schmidt P (2014) Using copulas to model time dependence in stochastic frontier models. Econom Rev 33(5–6):497–522

Amsler C, Prokhorov A, Schmidt P (2016) Endogeneity in stochastic frontier models. J Econom 190(2):280–288

Azzalini A (2005) The skew-normal distribution and related multivariate families. Scand J Stat 32(2):159–188

Bădin L, Simar L (2009) A bias-corrected nonparametric envelopment estimator of frontiers. Econometr Theor 25(5):1289–1318

Bartelsman EJ, Gray W (1996) The NBER manufacturing productivity database. NBER technical, Working Paper

Battese GE, Corra GS (1977) Estimation of a production frontier model: with application to the pastoral zone of eastern Autralia. Aus J Agr Resour Econ 21(3):169–179

Berndt ER, Hall BH, Hall RE, Hausman JA (1974) Estimation and inference in nonlinear structural models. Ann Econ Soc Meas 3(4):653–665

Burnham KP, Anderson DR (2004) Multimodel inference. Understanding AIC and BIC in model selection. Socio Meth Res 33(2):261–304

Carree M (2002) Technological inefficiency and the skewness of the error component in stochastic frontier analysis. Econ Lett 77:101–107

Carta A, Steel MFJ (2012) Modelling multi-output stochastic frontiers using copulas. Comput Stat Data Anal 56(11):3757–3773

Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE (2005) An introduction to efficiency and productivity analysis. Springer, New York, NY

Domma F (2004) Kurtosis diagram for the log-dagum distribution. Statistica & Applicazioni 2(2):3–23

Domma F, Perri P (2009) Some developments on the log-dagum distribution. Stat Method Appl 18:205–220

Feng Q, Horrace WC, Wu GL (2015) Wrong skewness and finite sample correction in parametric stochastic frontier models. center for policy research – The Maxwell School, working paper N. 154

Gómez-Déniz E, Pérez-Rodriguez JV (2015) Closed-form solution for a bivariate distribution in stochastic frontier models with dependent errors. J Prod Anal 43(2):215–223

Greene WH (1990) A gamma-distributed stochastic frontier model. J Econom 46(1-2):141–163

Green A, Mayes D (1991) Technical inefficiency in manufacturing industries. The Econ J 101:523–538

Hafner C, Manner H, Simar L (2016) The "wrong skewness" problem in stochastic frontier model: a new approach. Econometr Rev. doi:10.1080/07474938.2016.1140284

Huynh V-N, Kreinovich V, Sriboonchitta S (2014) Modeling dependence in econometrics. Springer, New York, NY

Joe H (1997) Multivariate models and dependence concepts. Chapman & Hall, London; New York

Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, 2nd edn, Vol. 2. Wiley, New York, NY

Judd KL (1998) Numerical methods in economics. The MIT, Cambridge

Kumbhakar SC, Lovell CAK (2000) Stochastic frontier analysis. Cambridge University Press, Cambridge

Lai H, Huang C (2013) Maximum likelihood estimation of seemingly unrelated stochastic frontier regressions. J Prod Anal 40(1):1–14

Lewis RA, McDonald JB (2013) Partially adaptive estimation of the censored regression model. Econometr Rev. doi:10.1080/07474938.2012.690691

Lin J-G, Xie F-C, Wei B-C (2013) Statistical diagnostics for skew-t-normal nonlinear models. Commun Stat-Simul C 38(10):2096–2110

Meeusen W, van den Broek J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Int Econ Rev 18(2):435–444

Meyer C (2013) The bivariate normal copula. Commun Stat-Theory Methods 42(13):2402–2422

Nelsen RB (1999) An introduction to Copula. Springer, New York, NY

Pal M, Sengupta A (1999) A model of FPF with correlated error components: an application to Indian agriculture. Sankhyā: The Indian Journal of Statistics, Series B (1960-2002) 61(2):337–350

Qian J, Sickles RC (2009) Stochastic frontiers with bounded inefficiency. Rice University, Working Paper

Shi P, Zhang W (2011) A copula regression model for estimating firm efficiency in the insurance industry. J Appl Stat 38(10):2271–2287

Simar L, Wilson P (2010) Inferences from cross-sectional, stochastic frontier models. Econom Rev 29(1):62–98

Simar L, Wilson P (2011) Estimation and inference in nonparametric frontier models: recent developments and perspectives. Found Trends Econometr 5(3–4):183–337

Smith MD (2008) Stochastic frontier models with dependent error components. Econom J 11:172–192

Stevenson RE (1980) Likelihood functions for generalized stochastic frontier estimation. J Econom 13:57–66

Tran KC, Tsionas EG (2015) Endogeneity in stochastic frontier models: Copula approach without external instruments. Econ Lett 133(C):85–88

Tsionas EG (2007) Efficiency measurement with the Weibull stochastic frontier. Oxf Bull Econ Stat 69(5):693–706

Wu L-C (2013) Variable selection in joint location and scale models of the skew-t-normal distribution. Commun Stat–Simul C 43(3):615–630

Zenga M (1985) Statistica descrittiva. Giappichelli Ed., Torino