The "wrong skewness" problem: a re-specification of Stochastic Frontiers.

Bonanno, Graziella and De Giovanni, Domenico and Domma, Filippo

Department of Economics, Statistics and Finance - University of Calabria, Department of Economics, Statistics and Finance - University of Calabria, Department of Economics, Statistics and Finance - University of Calabria

2015

The "wrong skewness" problem: a re-specification of Stochastic Frontiers.

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

Department of Economics, Statistics and Finance

University of Calabria - Italy.

[1]Corresponding author. Ponte Bucci Cubo 0C. 87036 Arcavacata di Rende (Cosenza) – ITALY. Tel: +39 0984492427. Fax: +39 0984492421. Email: f.domma@unical.it
Graziella Bonanno receives a Research Fellowship from the Regione Calabria and EU Commission. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of EU Commission and Regione Calabria.

The "wrong skewness" problem: a re-specification of Stochastic Frontiers.

Abstract

In this paper, we study the so-called "*wrong skewness*" anomaly in Stochastic Frontiers (SF), which consists in the observed difference between the expected and estimated sign of the asymmetry of the composite error. We propose a more general and flexible specification of the SF model, introducing dependence 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 in 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 displays wrong skewness, 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. In the second dataset, where the classic SF does not display any anomaly, estimation of our model provides evidence of the presence of both dependence between the error components and asymmetry of the random error.

Keywords: Stochastic frontier models, Skewness, Generalised Logistic distribution, Dependence, Copula functions.

JEL codes: C13, C18, C46, D24.

1 Introduction

The basic formulation of a production Stochastic Frontier (SF) model[2] can be expressed as $y = f(\mathbf{x};\boldsymbol{\beta})e^{\epsilon}$, where y is the firm production, $\mathbf{x}$ is a vector of inputs; $\boldsymbol{\beta}$ is vector of unknown parameters. The error term, $\epsilon = v - u$, is assumed to be made of two statistically independent components, a positive random variable, said u, and a symmetric random variable, said v. While u reflects the difference between the observed value of y and the frontier and it can be interpreted as a measure of firms' inefficiency, v captures random shocks, measurement errors and other statistical noise.

One major difficulty analysts often face when estimating a SF model is related to the choice of the distribution of random variables u and v. Different combinations have been proposed, including the normal-half normal model (Aigner et al., 1977), the normal-exponential model (Meeusen and van de Broek, 1977), normal–truncated normal model (Battese and Corra, 1977) and normal–gamma model Greene (1990). Perhaps the range of alternatives has been so far limited by computational challenges due to tractability issues of the convolution between the two error components. The choice of distributional specification is sometimes a matter of computational convenience.

The limited alternatives of possible distributions also poses empirical challenges. For instance, several authors have addressed the problem related to 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 ϵ is

$$E\left\{[\epsilon - E(\epsilon)]^3\right\} = -E\left\{[u - E(u)]^3\right\}, \tag{1}$$

thereby meaning, for example, a positive skewness for the inefficiency term u implies an expected negative skewness for the composite error ϵ. However, in many applications residuals display the wrong sign. This is called in literature the "*wrong skewness*" anomaly in *SF* models, initially highlighted by Green and Mayes (1991). To overcome this issue, several authors have proposed the use of distribution functions with negative asymmetry for inefficiency component. In partic-

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

ular, Carree (2002) uses the Binomial probability function, Tsionas (2007) suggests the Weibull distribution and Qian and Sickles (2009) whilst Almanidis and Sickles (2011) consider a double truncated Normal distribution.

More recent attempts to obtain the desired direction of residual skewness are Feng et al. (2013) where authors propose a finite sample adjustment to existing estimators and Hafner et al. (2013) where authors use an artificial truncation.

In this paper we argue that the wrong skewness problem has been only partially addressed because the relationship described by equation (1), and the consequent discussion about the wrong skewness anomaly, is a direct consequence of all the assumptions underlying the specification of the basic formulation of *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[3]

$$\begin{aligned} E\left\{[\epsilon - E(\epsilon)]^3\right\} &= -E\left\{[u - E(u)]^3\right\} + E\left\{[v - E(v)]^3\right\} \\ &+ 3cov\left(u^2, v\right) - 3cov\left(u, v^2\right) - 6\left[E(u) - E(v)\right] cov(u, v) \end{aligned} \quad (2)$$

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 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 the positive, negative dependence and the special case of independence according to the value of the 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

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

in Smith (2008), who uses FGM copulas to relax the assumption of independence between the two error terms. In a basic economic setting and with simple marginal distribution, Smith (2008) points out that the introduction of 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 the model with an exponential distribution for the inefficiency error, a logistic distribution for the random error. We propose a first generalization of Smith (2008) by using a Generalized Logistic (GL) distribution for the random error. This distribution describes situations of symmetry or asymmetry (positive or negative) according to values that takes on one of its parameters. This allows us to analyse the statistical properties of a model in which 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 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.

Our model allows for statistical dependence through copulas in a straightforward manner. It can be used to explicate the importance of including dependence in the economic context because it contributes to capture the effects of shocks that could affect both error components.

The 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 technical efficiency. The new specification of SF models is reported 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 two the applications; in particular, *Section 4.1* shows the estimations on data from NBER manufacturing productivity database that contains annual information on US manufacturing industries. We propose this example in order to verify our models in case of wrong skewness. In *Section 4.2* we test our tool on data from AIDA dataset including details of the Italian manufacturing firms. The implementation of traditional SF on this data does not imply wrong skewness. In *Section 5* we conclude. *Appendix A* presents the proof of our *Proposition 1* and *Appendix B* derives of Technical Efficiency scores. Despite the semi–

closed formula for the composite error function, estimation of our examples requires numerical discretization of the density. In this paper we use Gaussian quadratures, and the entire procedure is described in *Appendix C*.

2 Stochastic Frontiers and Copula functions

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

$$\overline{\mathbf{y}} = \overline{\mathbf{x}}\boldsymbol{\beta} + v - u \tag{3}$$

where **y** is a $(N \times 1)$ vector of firms' outputs; **x** is a $(N \times K)$ matrix of inputs; $\boldsymbol{\beta}$ is a $(K \times 1)$ vector of unknowns elasticities; **v** is a $(N \times 1)$ vector of random errors; **u** is a $(N \times 1)$ vector of random variables describing the inefficiencies associated to each firm (for a detailed discussion see Kumbhakar and Lovell, 2000).[4]

To complete the description of the model we need to specify the distributional properties of random variables (u, v). The standard specification assumes independence between the random error and the inefficiency error, and normal distribution for both random variables (though the inefficiency error must be truncated at zero to guarantee positiveness). We depart from this specification, by considering a general joint density $f_{u,v}(\cdot, \cdot, \Theta)$ 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 $\epsilon := v - u$ is obtained by convolution of two dependent random variables u and v, *i.e.*

$$f_\epsilon(\epsilon) = \int_{\Re^+} f_{u,v}(u, \epsilon + u) du \tag{4}$$

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

[4]Here and throughout the rest of the paper, overlined variables denote logarithmic transformation of original variables. For example $\overline{y} = log(y)$.

Copulas are widely appreciated tools used for the construction of joint distribution functions. To highlight the potential of this tool, it is sufficient to consider 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 stands in Sklar's theorem which proves how copulas link joint distribution functions to their one-dimensional margins. Indeed, according to Sklar's theorem any bivariate distribution $H(x, y)$ of 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, allow us to construct a joint distribution.

For the seek of parsimony, in this paper we do not include the rigorous construction of 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 five steps:

1. Choice of marginal distributions for the inefficiency error and the random error. We denote with $f_u(\cdot),\ g_v(\cdot)$ and $F_u(\cdot),\ G_v(\cdot)$ their probability density functions and distribution functions, respectively.

2. 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 $f(u, v)$ is given by 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 $c\big(F(u), G(v)\big) = \frac{\partial^2 C(F(u),G(v))}{\partial F(u)\partial G(v)}$ is the density copula.

4. The probability density function of the composite error $f_\epsilon(\cdot; \Theta)$ is obtained by convolution of the joint density as in (4). Now, observed that $\epsilon_i = \overline{y}_i - \overline{\mathbf{x}}_i\boldsymbol{\beta}$, the likelihood function is given by

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

being $\bar{\mathbf{x}}_i$ the i^{th} row of matrice $\bar{\mathbf{x}}$.

5. Finally, the Technical Efficiency (TE_Θ) is:

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

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 $C(\cdot,\cdot)$. 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 order to estimate the three components described in equation (2), which determine the sign of the asymmetry of the composite error, we must use a specification such that the shape of the *pdf* of v can be asymmetric (positive or negative) or symmetric according to the value of one of its parameters, a *pdf* for u with positive skewness in order to describe the specific characteristics of the distribution of the inefficiency and a dependence structure between u and v such that it can describe the situations of positive, negative dependence, or the particular case of independence.

To this end, we choose the Generalized Logistic (GL) distribution for the random error v, the Exponential distribution for the inefficiency error u and the FGM copula function for dependence structure between u and v. In table 1, we report the main features of these distributions.

[Table 1 about here.]

The parameter α_v of the GL distribution is an indicator of the direction of the skewness (the distribution is symmetric for $\alpha_v = 1$, asymmetric negative for $\alpha_v \in (0, 1)$ and asymmetric positive for $\alpha_v > 1$), while λ_v is the location parameter. The choice of the Generalized Logistic distribution makes our results directly comparable with those of Smith (2008), who uses a Standard Logistic distribution. Our results thus specialize to Smith (2008) with $\alpha_v = 1$ and $\lambda_v = 0$. Moreover, it is worth recalling that the FGM copula describes a situation of negative dependence, independence or positive dependence according to the parameter θ is less than, equal to or greater than zero, respectively.

The following *Proposition* reports the semi-explicit formulation for the *pdf* of the composite error in terms of linear combination of Hypergeometric functions[5], the expected value, the variance and the third central moment of the composite error.

Proposition 1 *Assuming that $u \sim Exp(\lambda_u)$, $v \sim GL(\lambda_v, \delta_v, \alpha_v)$ and the dependence between u and v is modeled by FGM copula. Let $k_1(\epsilon)$ be defined as $k_1(\epsilon) = \exp\{-\frac{\epsilon-\lambda_v}{\delta_v}\}$.*

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

$$\begin{aligned} f_\epsilon(\epsilon;\Theta) =& w_1(\epsilon)_2F_1\big(\alpha_v+1, \frac{\delta_v}{\delta_u}; \frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\big)+ \\ & w_2(\epsilon)_2F_1\big(\alpha_v+1, 2\frac{\delta_v}{\delta_u}+1; 2\frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\big)+ \\ & w_3(\epsilon)_2F_1\big(2\alpha_v+1, \frac{\delta_v}{\delta_u}+1; \frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\big)+ \\ & w_4(\epsilon)_2F_1\big(2\alpha_v+1, 2\frac{\delta_v}{\delta_u}+1; 2\frac{\delta_v}{\delta_u}+2; -k_1(\epsilon)\big) \end{aligned} \tag{8}$$

where the functions $w_1(.)$, $w_2(.)$, $w_3(.)$ and $w_4(.)$ are, respectively, defined as:

$$w_1(\epsilon) = (1-\theta)\frac{\alpha_v k_1(\epsilon)}{\delta_v+\delta_u} \qquad w_2(\epsilon) = 2\theta\frac{\alpha_v k_1(\epsilon)}{2\delta_v+\delta_u}$$
$$w_3(\epsilon) = 2\theta\frac{\alpha_v k_1(\epsilon)}{\delta_v+\delta_u} \qquad w_4(\epsilon) = -4\theta\frac{\alpha_v k_1(\epsilon)}{2\delta_v+\delta_u}$$

[5]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 $(d)_i = d(d+1)\dots(d+i-1)$ is the Pochhammer symbol, with $(d)_0 = 1$. In our case, $a = 2\alpha_v+1$, $b = 2\alpha_v + \frac{\delta_v}{\delta_u}$, $c = 2\alpha_v + \frac{\delta_v}{\delta_u} + 1$ and $s = -k_1^{-1}$.

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

$$E[\epsilon] = -\delta_u + \lambda_v + \delta_v[\Psi(\alpha_v) - \Psi(1)], \tag{9}$$

$$V[\epsilon] = \delta_u^2 + \delta_v^2[\Psi'(\alpha_v) + \Psi'(1)] - \theta\ \delta_u\delta_v\ [\Psi(2\alpha_v) - \Psi(\alpha_v)] \tag{10}$$

and

$$\begin{aligned} E[\epsilon - E(\epsilon)]^3 = & -2\delta_u^3 + \delta_v^3[\Psi''(\alpha_v) - \Psi''(1)] + \frac{3}{2}\theta\delta_u\{-\delta_v^2[\Psi'(2\alpha_v) - \Psi'(\alpha_v)] + \\ & \delta_u\delta_v[\Psi(2\alpha_v) - \Psi(\alpha_v)] - [\lambda_v + \delta_v(\Psi(2\alpha_v) - \Psi(1))]^2 + \\ & [\lambda_v + \delta_v(\Psi(\alpha_v) - \Psi(1))][\lambda_v + \delta_v(2\Psi(2\alpha_v) - \Psi(\alpha_v) - \Psi(1))]\} \end{aligned} \tag{11}$$

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

Proof. See *Appendix A*.

To appreciate the flexibility of our model, we point out that according 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 random error (via parameter α_v) and of the dependence (via parameter θ) between u and v on the variance of composite error.

In particular, we compare four variances of the composite error corresponding to 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 find the special case

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

which overlaps Smith (2008) in the case of symmetry of v and dependence between u and v (it corresponds to variance of ϵ of model (D,S)). Moreover, to make simple discussion, we highlight that the variance of composite error in the cases of (a) independence and asymmetry and (b) independence and symmetry, are given, respectively, by $V_\epsilon^{(I,A)} = \delta_u^2 + \delta_v^2\left[\psi'(\alpha_v) + \psi'(1)\right]$ and $V\epsilon^{(I,S)} = \delta_u^2 + \frac{\pi^2}{3}\delta_v^2$. Obviously, the variance of composite error in the case of dependence and asymmetry is $V\epsilon^{(D,A)} = V(\epsilon)$ reported in eq. (10).

Figure 1 plots the variance of ϵ as 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.

[Figure 1 about here.]

[Figure 2 about here.]

Next, we show the effects of α_v on the distribution function of ϵ.[6] In fact, figure 2 shows how the asymmetry of random error affects the distribution of the composite error. Imposing maximum positive dependence between u and v ($\theta = 1$), we plot different density functions for different values of α_v and observe that α_v impacts not only on the shape of the density, but also, and more importantly, on the behavior of the distribution at the tails. The effect is more pronounced in case of negatively skewed distributions of random error. This finding explains the impact on the variance observed above: negative skewness assigns much more probability mass to extreme negative values of ϵ than positive skewness.

[6] We have analysed the impact of the dependence structure on the density of ϵ. Smith (2008) show this effect in the case of *symmetric-v*. We observe the same results in different conditions of skewness for v (negative or positive). For this reason we do not report here the plots.

The empirical literature often faces estimated skewed density functions of the composite error contrasting theoretical predictions of the model (*wrong skewness* anomaly). In this respect, there is no general consensus on the interpretation of this misalignment between assumptions and observed facts. For instance, Kumbhakar and Lovell (2000) ascribe the misalignment to economically significant model misspecifications, while Smith (2008) argues that the observed skewness may arise from the dependence between the random error and inefficiency. Here, we contribute to the debate by suggesting one more possible explanation: it would be the interaction between the dependence (as argued by Smith) and the fundamental asymmetry of the distribution of random error.

4 Empirical examples of production frontiers

In what follows, we report two examples.[7] We use two different data samples, one in which a case of "*wrong skewness*" occurs, and one in which it does not occur.

4.1 Wrong skewness in data from NBER database

We test our model for a SF production frontier using data from NBER manufacturing productivity database (Bartelsman and Gray, 1996). This archive is free available online and contains annual information on US manufacturing industries since 1958 to the present. We focus on data of 1979 since, after checking for the asymmetry of OLS residuals, we find the presence of strong positive skewness in 1979, while negative skewness was expected from the traditional model (this is also showed in Hafner et al. (2013)). The case of wrong skewness is confirmed when the classic production SF is estimated.[8]

In the underlying economic model, the variable value added is our output and total employment ($lemployment$) and capital stock ($lcap$) are input factors (all variables are in logs). The frontier assumes the Cobb-Douglas functional form. We want to highlight that our specification of the

[7]The maximisation routine is been developed in the software R-projet using "maxLik" package and then the estimates are been controlled with the algorithm discussed in Appendix C.

[8]When positive skewness is found, classic SF estimates coincide with the OLS, because we reject the hypotheses of presence of inefficiency.

random error through a GL distribution includes a location parameter, λ_v that acts as intercept in the regression. Therefore, to avoid identification issues, our regression model does not include intercept (see table 2 for more details).[9]

[Table 2 about here.]

We report the results in table 3 where significant coefficients are in bold (t-statistics are reported in parenthesis).

The last two columns report the *classic* SF estimates (one model without the intercept, one with the intercept in the production function) where the residuals are assumed to be normally distributed and there is independence between u and v. The results show that SF estimates coincide with OLS because we reject the hypotheses of presence of inefficiency. The estimated measure of the contribute of the variance of u to the total variance is very close to 0. Standard residual analysis shows that the model is not correctly specified.[10] For these reasons, we do not comment results about the *classic* SF estimates.

Turning to our models, the attention goes first at the parameters of marginal distributions and association measure θ that is not statistically significant. All the other estimated parameters are widely significant, except the location parameter of the Exponential assigned to inefficiency error, that is not statistically different from 0 for all specifications.

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 -22.44, followed by IS (-21.15), DA (-20.44) and DS (-19.19).[11] Following Burnham and Anderson (2004), even if each of our four models may be indifferent to others, 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 random error and not involving dependence structures.

---

[9] In both the empirical applications, however, the estimates of λ_v are very close to the estimate of the intercepts of the classic SFs (see table 3 in this sub-section about NBER data and table 6 in the next sub-section 4.2 about AIDA data).

[10] Also the software R-project provides the following warning message: "The residuals of the OLS estimates are right-skewed; this might indicate that there is no inefficiency or that the model is misspecified.".

[11] Burnham and Anderson (2004) consider the measure $\Delta_i = AIC_i - AIC_{min}$. According with the authors, models having $\Delta_i \leq 2$ have substantial evidence, those for which $4 \leq \Delta_i \leq 7$ have less support, and models having $\Delta_i > 10$ have no support.

[Table 3 about here.]

Table 4 contains some descriptive statistics about the estimated parameters and the composite error ϵ. Each column represents one model, whose statistical characteristics are in table 2. It is worth noticing that the true direction of asymmetry is measured as factorisation of the sum of deviations from the median (Zenga, 1985).[12] One element of this decomposition is $E[\epsilon - E(\epsilon)]^3$, that is the measure derived in equation (11). In table 4 we report the contributes of single components to explain $E[\epsilon - E(\epsilon)]^3$ and $E[\epsilon - Me(\epsilon)]^3$.

We find positive skewness of ϵ in IA and DA models, in which the v-component is strongly positive, while DS models show wrong skewness. For IS model, we can accept the symmetry of ϵ (all the skewness measures are very close to 0). In fact, the sign of $E[\epsilon - Me(\epsilon)]^3$ and $\sum[\hat{\epsilon} - Me(\hat{\epsilon})]^3$ is the same for IA and DA models, it is opposite for DS. We remark that: *i*) IS assumes *a priori* that v-component and dependence-component are equal to 0, as in classic SF; *ii*) dependence-component is negative for both models with dependence structure, DA and DS, but dependence is statistically rejected in this data sample.

[Table 4 about here.]

Finally, in table 5 we report some descriptive statistics on estimated Technical Efficiency (TE) for each model. [13] In particular, the bias evident for classic SFs and IS model is solved in our preferred model (IA).

[Table 5 about here.]

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

$$E[\epsilon - Me(\epsilon)]^3 = E[\epsilon - E(\epsilon)]^3 + [E(\epsilon) - Me(\epsilon)]^3 + 3[E(\epsilon) - Me(\epsilon)]V(\epsilon). \quad (13)$$

[13]The derivation of TE_{Θ} scores is reported in *Appendix B*.

4.2 Application on a sample of Italian manufacturing firms

We use data from AIDA ("Analisi Informatizzata delle Aziende Italiane"), that is a database containing financial and accounting information of Italian companies.

We use again a Cobb-Douglas production function where the dependent variable is the value added representing the firms' output, while labour and capital are the traditional inputs. Moreover, we introduce ICT and R&D investments as additional inputs. All variable, referring to 2009, are in logs.[14]

[Table 6 about here.]

Going to examine the results from table 6, we highlight the robustness of the estimates across our models and the significance of all fitted parameters (t-statistics are in bracket). Moreover, from AIC measure, 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 better fit is due to the more general DA model (AIC 1089.84), while the worst is the more parsimonious IS (1112.18). The results highlight the presence of positive dependence (θ is equal to 0.7016 in DA) in this data sample.

Switching to analyses the descriptive statistics of the various models (table 7), there is not case of wrong skewness in the simpler model IS.[15]

[Table 7 about here.]

[14] We calculate ICT and R&D investments as percentage of yearly sales. This percentage is from EFIGE dataset ("European Firms in a Global Economy: internal policies for external competitiveness"), which combines measures of firms' international activities with quantitative and qualitative information with focus on R&D and innovation.

[15] We propose this example also to show the validity of our models in case of absence of wrong skewness.

5 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 that appear to be overly restrictive. In fact, relaxing the hypotheses of symmetry of the random error and independence between the components of the composite error, we obtain a re-specification of Stochastic Frontiers sufficiently flexible that allows us to explain the difference between the expected and the estimated sign of the asymmetry of the composite error, found in various applications of the classic Stochastic Frontier.
The decomposition the third moment of the composite error in 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 enables us to reinterpret the unusual asymmetry in the composite error by measuring the contribution of each component in the model. This is shown in one of the two empirical examples, i.e. on data from NBER archive, for which a case of wrong skewness is reported (present) with the *classic* SF specification.

When wrong skewness occurs, estimations with classic SF correspond to OLS estimations, and the inefficiency scores are zeros. This misleads to the conclusion of absence of inefficiency. Our specification allows to overcome this difficulties, as witnessed in both empirical applications, where our estimation of the output elasticities with respect to inputs are quite robust against to the standard SF specification, but estimated efficiency scores are lower than the unity.

References

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

P. Almanidis and R.C. Sickles. *Exploring Research Frontiers in Contemporary Statistics and Econometrics: A Festschrift in Honor of Leopold Simar*, chapter Skewness problem in Stochastic Frontier Models: Fact or Fiction? New York: Springer, 2011.

E. J. Bartelsman and W. Gray. The NBER manufacturing productivity database. *NBER technical, Working Paper*, 1996.

G. E. Battese and G.S. Corra. Estimation of a production frontier model: With application to the pastoral zone of eastern australia. *Australian Journal of Agricultural Economics*, 21:169–179, 1977.

G. E. Battese, T. J. Coelli, D. S. P. Rao, and C. J. OD́onnell. *An Introduction to Efficiency and Productivity Analysis*. 2005.

E. Berndt, B. Hall, R. Hall, and J. Hausman. Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 4:653–665, 1974.

K.P. Burnham and D.R. Anderson. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods Research*, 2004.

M.A. Carree. Technological inefficiency and the skewness of the error component in stochastic frontier analysis. *Economics Letters*, 77:101–107, 2002.

Q. Feng, W.C. Horrace, and G.L. Wu. Wrong skewness and finite sample correction in parametric stochastic frontier models. 2013.

A. Green and D. Mayes. Technical inefficiency in manufacturing industries. *Economic Journal*, 101:523538, 1991.

W. H. Greene. A gamma-distributed stochastic frontier model. *Journal of Econometrics*, 46: 141–163, 1990.

C.M. Hafner, H. Manner, and L. Simar. The ”wrong skewness” problem in stochastic frontier model: a new approach”. *Insitutute of Statistics, Biostatistics and Actuarial Sciences*, 2013.

Kenneth L. Judd. *Numerical Methods in Economics*. The MIT Press, 1998.

S. C. Kumbhakar and C. A. K. Lovell. *Stochastic Frontier Analysis*. Cambridge University Press, 2000.

D. Meeusen and J. van de Broek. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2):435–444, 1977.

R. B. Nelsen. *An Introduction to Copula*. Springer, 1999.

J. Qian and R. C. Sickles. Stochastic frontiers with bounded inefficiency. *Rice University, WP*, 2009.

M. D. Smith. Stochastic frontier models with dependent error components. *Econometrics Journal*, 11:172–192, 2008.

E. G. Tsionas. Efficiency measurement with the weibull stochastic frontier. *Oxford Bulletin of Economics and Statistics*, 69:693–706, 2007.

M. Zenga. *Statistica descrittiva*. Giappichelli, 1985.

A Proof of Proposition 1

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

Lemma 2 *1. If $U \sim Exp(\delta_u)$ then*

- *r-th moment is $E(U^r) = \lambda_u^r \Gamma(r+1)$. 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 $E\left[U^r F(U)\right] = E(U^r)\left(1 - \frac{1}{2^{r+1}}\right)$*

2. If $V \sim GL(\lambda_v, \delta_v, \alpha_v)$ then

- $E(V) = \lambda_v + \delta_v\left[\Psi(\alpha_v) - \Psi(1)\right]$;
- $E(V^2) = \delta_v^2\left[\Psi'(\alpha_v) + \Psi'(1)\right] + \left[\lambda_v + \delta_v\left(\Psi(\alpha_v) - \Psi(1)\right)\right]^2$;
- $E(V^3) = \delta_v^3\left[\Psi''(\alpha_v) - \Psi''(1)\right] + \left[\lambda_v + \delta_v\left(\Psi(\alpha_v) - \Psi(1)\right)\right]^3 + 3\delta_v^2\left[\lambda_v + \delta_v\left(\Psi(\alpha_v) - \Psi(1)\right)\right]\left[\Psi'(\alpha_v) + \Psi'(1)\right]$
- *Denoted with $G(v) = \left(1 + e^{-\frac{v-\lambda_v}{\delta_v}}\right)^{-\alpha_v}$ the distribution function of the random variable V, we have $E\left[V^k G(V)\right] = \frac{1}{2}E\left[V^k|2\alpha_v, \lambda_v, \delta_v\right]$, where $E\left[.|2\alpha_v, \lambda_v, \delta_v\right]$ is the expectation with respect to the GL with parameters $2\alpha_v$, λ_v and δ_v.*

3. if $(U, V) \sim f_{u,v}(u, v) = f(u)g(v)\left[1 + \theta\left(1 - 2F(u)\right)\left(1 - 2G(v)\right)\right]$ then

$$
\begin{aligned}
E\left(U^r V^k\right) &= (1+\theta)E(U^r)E(V^k) - 2\theta\left\{E(U^r)E\left[V^k G(V)\right] + E\left[U^r F(U)\right]E(V^k) - \right.\\
&\quad \left. 2E\left[U^r F(U)\right]E\left[V^k G(V)\right]\right\} = \\
&= E(U^r)E(V^k) + \theta\left(\frac{1}{2^r} - 1\right)E(U^r)\left\{E(V^k) - E(V^k|2\alpha_v, \lambda_v, \delta_v)\right\}
\end{aligned}
$$

Now, we can prove the *Proposition* 1.

1. The *pdf* of composite error is $f(\epsilon) = \int_{\Re^+} f(u, \epsilon + u)du$ where $f(u, \epsilon + u) = f(u)g(\epsilon + u)c\left(F(u), G(\epsilon + u)\right)$.

Given that $c(\cdot,\cdot)$ is a density copula of a FGM copula, we have

$$\begin{aligned} f(u,\epsilon+u) &= (1+\theta)f(u)g(\epsilon+u) - 2\theta f(u)g(\epsilon+u)G(\epsilon+u) \\ &- 2\theta f(u)g(\epsilon+u)F(u) + 4\theta f(u)g(\epsilon+u)F(u)G(\epsilon+u) \qquad \text{(A.1)} \end{aligned}$$

Using (A.1), we have $f(\epsilon) = (1+\theta)I_1 - 2\theta\left\{I_2 + I_3 - 2I\right\}$, where $I = \int_{\Re^+} f(u)g(\epsilon+u)F(u)G(\epsilon+u)du$, and I_i, for $i = 1,2,3$ are special cases of I.

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

$$f(u)g(\epsilon+u)F(u)G(\epsilon+u) = \frac{\alpha_v k_1(\epsilon)}{\delta_u\delta_v} e^{-\frac{u}{\delta_u}-\frac{u}{\delta_v}}\left(1-e^{-\frac{u}{\delta_u}}\right)\left(1+k_1(\epsilon)e^{-\frac{u}{\delta_u}}\right)^{-2\alpha_v-1} \text{(A.2)}$$

where $k_1(\epsilon) = e^{-\frac{\epsilon-\lambda_v}{\delta_v}}$. After algebra, we can write

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

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

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

Bearing in mind that for hypergeometric function is true the following

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

We obtain

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

2. By Lemma, we can to verify that

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

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

  and

  $$
  \begin{aligned}
  cov(U, V^2) &= -\frac{\theta}{2}\delta_u \left\{ \delta_v^2 \left[\Psi'(\alpha_v) - \Psi'(2\alpha_v)\right] + \left[\lambda_v + \delta_v\left(\Psi(\alpha_v) - \Psi(1)\right)\right]^2 + \right. \\
  &\left. - \left[\lambda_v + \delta_v\left(\Psi(2\alpha_v) - \Psi(1)\right)\right]^2 \right\}
  \end{aligned}
  $$

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

B Calculation of TE scores

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

We can write

$$\begin{aligned} TE_\Theta =& E[e^{-u}|\epsilon = \epsilon^*] = \frac{1}{f_\epsilon(.;\Theta)} \int_{\Re+} e^{-u} f_{u,v}(u, x+u;\Theta) du \\ =& \frac{1}{f_\epsilon(.;\Theta)} \int_{\Re+} e^{-u} f(u, \epsilon+u) du = \frac{1}{f_\epsilon(.;\Theta)} \int_{\Re+} e^{-u} f(u, \epsilon+u) du \end{aligned} \tag{B.1}$$

where $f(u, \epsilon + u)$ is derived in equation A.1.

After algebra, we obtain:

$$TE_\Theta = E[e^{-u}|\epsilon] = \frac{\overline{\omega}_1(\epsilon)\overline{H}_1(\epsilon) + \theta[\overline{\omega}_1(\epsilon)\overline{H}_1(\epsilon) - 2\overline{\omega}_2(\epsilon)\overline{H}_2(\epsilon) - 2\overline{\omega}_3(\epsilon)\overline{H}_3(\epsilon) + 4\overline{\omega}_4(\epsilon)\overline{H}_4(\epsilon)]}{\omega_1(\epsilon)H_1(\epsilon) - \theta[\omega_1(\epsilon)H_1(\epsilon) - 2\omega_2(\epsilon)H_2(\epsilon) - 2\omega_3(\epsilon)H_3(\epsilon) + 4\omega_4(\epsilon)H_4(\epsilon)]} \tag{B.2}$$

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

$$
\begin{aligned}
\overline{H}_1 &= {}_2F_1\left(\alpha_v + 1, \frac{\delta_v}{\delta_u} + \delta_v + 1; \frac{\delta_v}{\delta_u} + \delta_v + 2; -k_1(\epsilon)\right) \\
H_1 &= {}_2F_1\left(\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right) \\
\overline{H}_2 &= \frac{1}{\left[\delta_v\left(\frac{1}{\delta_u} + 1\right)\right]} {}_2F_1\left(\alpha_v + 1, \delta_v\left(\frac{1}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{1}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right) \\
&\quad - \frac{1}{\left[\delta_v\left(\frac{2}{\delta_u} + 1\right)\right]} {}_2F_1\left(\alpha_v + 1, \delta_v\left(\frac{2}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{2}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right) \\
H_2 &= {}_2F_1\left(\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right) \\
\overline{H}_3 &= {}_2F_1\left(2\alpha_v + 1, \frac{\delta_v}{\delta_u} + \delta_v + 1; \frac{\delta_v}{\delta_u} + \delta_v + 2; -k_1(\epsilon)\right) \\
H_3 &= {}_2F_1\left(2\alpha_v + 1, \frac{\delta_v}{\delta_u} + 1; \frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right) \\
\overline{H}_4 &= \frac{1}{\left[\delta_v\left(\frac{1}{\delta_u} + 1\right)\right]} {}_2F_1\left(2\alpha_v + 1, \delta_v\left(\frac{1}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{1}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right) \\
&\quad - \frac{1}{\left[\delta_v\left(\frac{2}{\delta_u} + 1\right)\right]} {}_2F_1\left(2\alpha_v + 1, \delta_v\left(\frac{2}{\delta_u} + 1\right) + 1; \delta_v\left(\frac{2}{\delta_u} + 1\right) + 2; -k_1(\epsilon)\right) \\
H_4 &= {}_2F_1\left(2\alpha_v + 1, 2\frac{\delta_v}{\delta_u} + 1; 2\frac{\delta_v}{\delta_u} + 2; -k_1(\epsilon)\right)
\end{aligned}
$$

and where the $\omega - functions$ are respectively defined as:

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

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

C The numerical procedure

The estimation of models like that described in Section 2 requires the ability to compute the density of composite error. Closed–form expressions for this quantity are available only in some few special cases such as the notable case addressed by Smith (2008). While in the previous section we provide one more example of 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 SF models.

Our approach is fairly simple. We approximate the convolution between u and v by means of numerical quadratures. To be more precise, set $\varepsilon = v - u$, its density function, $f_\varepsilon(\cdot; \Theta)$, is obtained by the convolution of of u and v:

$$f_\varepsilon(x; \Theta) = \int_0^\infty f_{u,v}(u, x + u; \Theta) du \tag{C.3}$$

Explicit evaluation of the integral in (C.3) is in general infeasible, keeping a potential range of possible joint densities almost unexplored. However, approximation of (C.3) by Gauss–Laguerre quadrature has proved to be easy and effective, and is reported below.

Let us first rewrite (C.3) as

$$f_\varepsilon(x; \Theta) = \int_0^\infty e^{-u} g_x(u) du, \tag{C.4}$$

with $g_x(u) = e^u f_{u,v}(u, x + u)$. Fix an integer m that we refer to as the order of quadrature and, 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 defined by the following system of linear equations[16, 17]

$$\int_0^\infty s^k e^{-s} ds = \sum_{h=1}^{n} \omega_h t_h^k \quad k = 1, \ldots, 2m - 1. \tag{C.5}$$

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

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

Then, we can write

$$f_\varepsilon(x;\Theta) \approx \sum_{h=1}^{m} \omega_h g_x(t_h). \tag{C.6}$$

As far as the function $g_x(\cdot)$ is Reimann–integrable over the interval $[0, \infty)$, standard results in numerical analysis ensure the goodness of the approximation.

We can thus approximate the integral appearing in (C.3) (and its gradient with respect to Θ) with a finite sum, and insert the approximated density function and its gradient into a Quasi–Newton–like iteration (however, from experience with the Normal/Half–Normal with FGM copula model, 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 with FGM copula case 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.

Figure 1: Plot of $VAR(\epsilon)$ for a production frontier with $\theta = -1$, $\theta = 0$ and $\theta = 1$ (α_v ranges between 0 and 2).

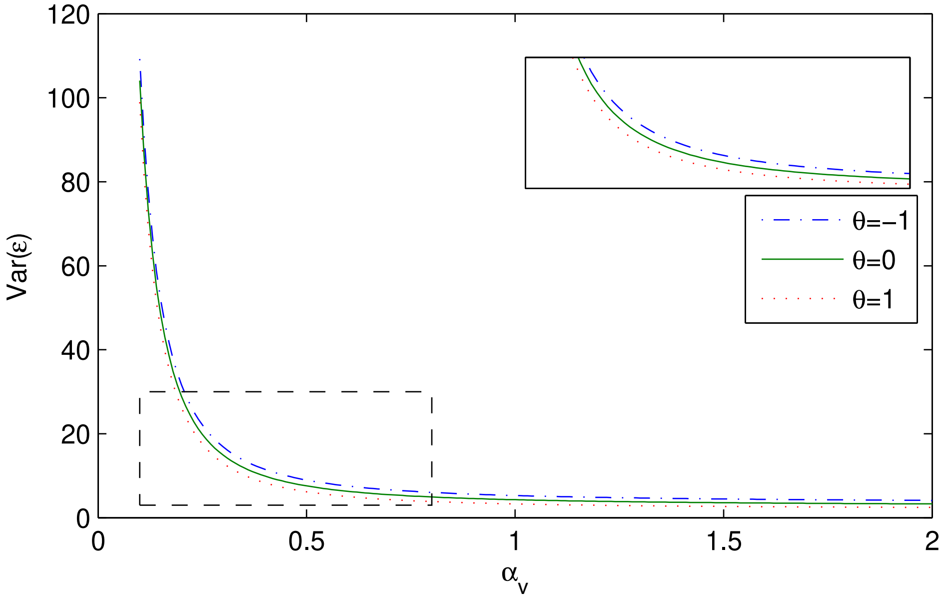

Figure 2: Density function of ϵ of a production frontier with $\lambda_v = 0$, $\delta_u = \delta_v = 1$, $\alpha_u = 1$ and $\theta = 1$ (α_v ranges between 0.25 and 3).

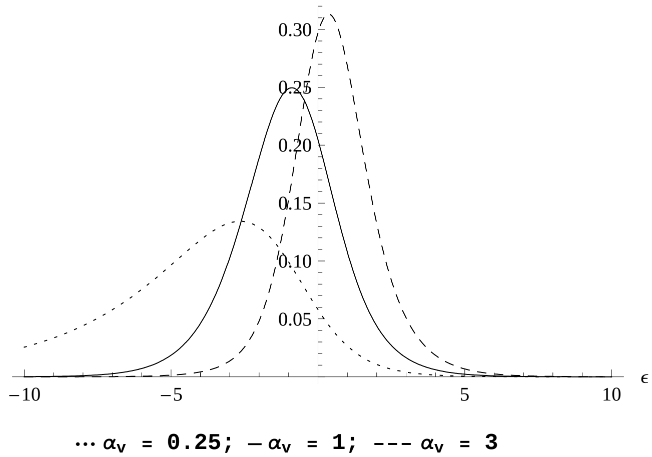

Table 1: Marginal distribution functions and FGM copula.

	Parameters	Density	Distribution
Exponential	$\delta_u > 0$	$\frac{1}{\delta_u} e^{-\frac{u}{\delta_u}}$	$1 - e^{-\frac{u}{\delta_u}}$
GL	$\alpha_v,\ \delta_v > 0, \lambda_v \in \Re$	$\frac{\alpha_v}{\delta_v} \frac{e^{-\frac{v-\lambda_v}{\delta_v}}}{\left(1+e^{-\frac{v-\lambda_v}{\delta_v}}\right)^{\alpha_v+1}}$	$(1 + e^{-\frac{v-\lambda_v}{\delta_v}})^{-\alpha_v}$
FGM	$\theta \in (-1, 1)$	$1 + \theta(1 - 2F_u)(1 - 2G_v)$	$F_u G_v \big(1 + \theta(1 - F_u)(1 - G_v)\big)$

Table 2: Summary of the statistical models.

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

Legend: IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; IS stands for independence and symmetry; finally, DS is the model with FGM dependence and symmetry.

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

	IA	DA	IS	DS	*Classic SF*	*Classic* SF Intercept
β_0						**2.4507**
						(8.85)
$\beta_1(lemployment)$	**0.8316**	**0.8179**	**0.7968**	**0.7995**	0.7373	**0.7920**
	(8.43)	(25.19)	(23.84)	(23.09)	(0.72)	(20.67)
$\beta_2(lcap)$	**0.1817**	**0.1864**	**0.1912**	**0.1879**	0.6097	**0.1913**
	(1.60)	(8.29)	(7.12)	(6.89)	(1.01)	(6.28)
δ_u	0.0132	0.0458	0.0002	0.0750		
	(0.54)	(0.79)	(0.01)	(0.79)		
α_v	**5.5389**	**2.6832**	-	**-**		
	(79.34)	(4.53)				
δ_v	**0.1361**	**0.1291**	**0.0994**	**0.1035**		
	(8.88)	(2.36)	(8.63)	(3.48)		
λ_v	**2.0938**	**2.2616**	**2.4218**	**2.5075**		
	(5.86)	(24.06)	(19.25)	(16.40)		
θ	-	0.9995	-	0.99998		
		(0.18)		(0.35)		
Obs	54	54	54	54	54	54
log-likelihood	17.22	17.22	15.58	15.59	-41.07	13.58
AIC	-22.44	-20.44	-21.15	-19.19	90.14	-17.17

Source: our elaborations on data from the NBER productivity database. The dependent variable is the value added (in log).

Legend: IA = model with independence and asymmetry; DA = model with FGM dependence and asymmetry; IS = model with independence and symmetry; DS = model with FGM dependence and symmetry.

Table 4: Summary measures and skewness of the composite error (data from NBER).

	IA	DA	IS	DS
δ_u	0.0132	0.0458	0.0002	0.0750
α_v	5.5389	2.6832	1	1
δ_v	0.1361	0.1291	0.0994	0.1035
λ_v	2.0938	2.2616	2.4218	2.5075
θ	0	0.9995	0	0.99998
$E(\epsilon)$	2.3794	2.3923	2.4216	2.4325
$V(\epsilon)$	0.0343	0.0323	0.0325	0.0331
$Me(\epsilon)$	2.3549	2.3741	2.4216	2.4331
$E[\epsilon - E(\epsilon)]^3$	0.005953	0.004435	-1.82E-11	-0.002381
u-component	*-0.000005*	*-0.000192*	*-1.82E-11*	*-0.000844*
v-component	*0.005957*	*0.004746*	*0*	*0*
dependence-component	*0*	*-0.000119*	*0*	*-0.001537*
$E[\epsilon - Me(\epsilon)]^3$	0.0085	0.0062	6.80E-09	-0.0024
$\sum[\hat{\epsilon} - Me(\hat{\epsilon})]^3$	0.5749	0.5125	0.4513	0.4498

Source: our elaborations on data from the NBER productivity database.

Legend1 IA = model with independence and asymmetry; DA = model with FGM dependence and asymmetry; IS = model with independence and symmetry; DS = model with FGM dependence and symmetry.

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

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

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

Table 5: Some descriptive statistics of Technical Efficiency (data from NBER).

	IA	DA	IS	DS	*Classic* SF	*Classic* SF - Intercept
Mean	0.9870	0.9563	0.9998	0.9305	1	0.999546
Stand. Dev.	0.0011	0.0028	2.26e-07	0.0085	0	6.18e-07
Min	0.9831	0.9492	0.9997	0.9066	1	0.999545
Max	0.9881	0.9587	0.9998	0.9385	1	0.999548

Source: our elaborations on data from the NBER productivity database.
Legend1 IA = model with independence and asymmetry; DA = model with FGM dependence and asymmetry; IS = model with independence and symmetry; DS = model with FGM dependence and symmetry.

Table 6: Estimations of production SF for the Italian manufacturing firms using data from AIDA (2009).

	IA	DA	IS	DS	*Classic* SF	*Classic* SF - Intercept
β_0						**2.5619**
						(20.78)
$\beta_1(llabour)$	**0.7497**	**0.7525**	**0.7372**	**0.7219**	**0.5583**	**0.7203**
	(37.39)	(37.78)	(36.62)	(35.92)	(20.73)	(32.59)
$\beta_2(lcapital)$	**0.0972**	**0.0970**	**0.0969**	**0.0979**	**0.1132**	**0.0970**
	(8.12)	(8.15)	(7.97)	(8.06)	(5.47)	(7.28)
$\beta_3(lICT)$	**0.0923**	**0.0909**	**0.1058**	**0.1188**	**0.2448**	**0.1110**
	(6.62)	(6.52)	(7.49)	(8.34)	(13.59)	(7.17)
$\beta_4(lR\&D)$	**0.0956**	**0.0947**	**0.1047**	**0.1054**	**0.2059**	**0.0986**
	(7.33)	(7.32)	(7.92)	(8.01)	(11.82)	(6.62)
δ_u	**0.3123**	**0.3499**	**0.2533**	**0.2688**		
	(16.12)	(9.85)	(10.07)	(6.01)		
α_v	**3.4756**	**2.4054**	-	-		
	(2.65)	(3.32)				
δ_v	**0.2205**	**0.2362**	**0.1939**	**0.2188**		
	(19.57)	(12.21)	(20.49)	(17.80)		
λ_v	**2.2115**	**2.3439**	**2.3254**	**2.2408**		
	(15.34)	(16.87)	(20.58)	(18.65)		
θ	-	**0.7016**	-	**0.7854**		
		(1.61)		(2.91)		
Mean Efficiency	0.7651	0.7432	0.7997	0.7887	0.99996	0.6897
Obs	939	939	939	939	939	939
log-likelihood	-538.56	-535.92	-549.09	-546.52	-751.81	-599.94
AIC	1093.12	1089.84	1112.18	1109.04	1515.62	1211.88

Source: our elaborations on data from the AIDA dataset. The dependent variable is the value added (in log).
Legend: IA = model with independence and asymmetry; DA = model with FGM dependence and asymmetry; IS = model with independence and symmetry; DS = model with FGM dependence and symmetry.

Table 7: Summary measures and skewness of the composite error (data from AIDA).

	IA	DA	IS	DS
δ_u	0.3123	0.3499	0.2533	0.2688
α_v	3.4756	2.4054	1	1
δ_v	0.2205	0.2362	0.1939	0.2188
λ_v	2.2115	2.3439	2.3254	2.2408
θ	0	0.7016	0	0.7854
$E(\epsilon)$	2.2678	2.2852	2.0722	1.9720
$V(\epsilon)$	0.1937	0.1961	0.1878	0.1836
$Me(\epsilon)$	2.2887	2.3057	2.0988	1.9890
$E[\epsilon - E(\epsilon)]^3$	-0.0364	-0.0404	-0.0325	-0.0506
u-component	*-0.0609*	*-0.0857*	*-0.0325*	*-0.0389*
v-component	*0.0246*	*0.0283*	*0*	*0*
dependence-component	*0*	*0.0170*	*0*	*-0.0117*
$E[\epsilon - Me(\epsilon)]^3$	-0.0485	-0.0525	-0.0475	-0.0599
$\sum[\hat{\epsilon} - Me(\hat{\epsilon})]^3$	-22.74	-24.25	-23.49	-22.37

Source: our elaborations on data from AIDA dataset.

Legend1 IA = model with independence and asymmetry; DA = model with FGM dependence and asymmetry; IS = model with independence and symmetry; DS = model with FGM dependence and symmetry.

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

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

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