Economics Bulletin

Volume 38, Issue 4

Cost Efficiency, Asymmetry and Dependence in US electricity industry

Graziella Bonanno
University of Trieste

Filippo Domma
University of Calabria

Abstract

We propose an empirical application of models derived in Bonanno et al. (2017) for estimating cost efficiency (CE) on data used by Greene (1990) to test Gamma distribution for the inefficiency component and by Smith (2008) to test the dependence between the two error terms of a Stochastic Frontier (SF). We also derive the closed–form of denisty function of the overall error term and the formula to calculate the Cost Efficiency (CE) scores.

Citation: Graziella Bonanno and Filippo Domma, (2018) "Cost Efficiency, Asymmetry and Dependence in US electricity industry", *Economics Bulletin*, Volume 38, Issue 4, pages 2379-2388
Contact: Graziella Bonanno - gbonanno@units.it, Filippo Domma - filippo.domma@unical.it.
Submitted: October 02, 2017. **Published:** December 10, 2018.

1 Introduction

Our aim, in this paper, is to show the effects of (*i*) the asymmetry of the random error term and (*ii*) the dependence between the two error terms of SF (random error term and the inefficiency component) in a cost function, by using the generalization of stochastic frontiers (SF) proposed by Bonanno et al. (2017) for production functions. We employ a well-known dataset used in previous literature in order to test some methodological improvements in the SF approach.

The basic formulation of a cost frontier model can be expressed as $c = f(\mathbf{Q}, \mathbf{P}; \boldsymbol{\beta})e^{\epsilon}$, where c is the firm-specific total costs, $\mathbf{Q}$ is a vector of outputs, $\mathbf{P}$ is a vector of inputs prices and $\boldsymbol{\beta}$ is the vector of unknown parameters (details are in Kumbhakar and Lovell, 2000). In the traditional SF specifications, the error term, ϵ, is assumed to be made up of two statistically independent components - a positive random variable, u, and a symmetric random variable, v. While u reflects the difference between the observed value of c and the frontier and it can be interpreted as a measure of firms' inefficiency, v captures random shocks, measurement errors and others statistical noise. We have that $\epsilon = u + v$ in a cost function. The existence of v implies random variations from the best frontier across firms.

We propose a model can capture the dependence structure between u and v, modeling it with a copula function that allows us to specify the joint distribution with different marginal probability density functions in a simple way. In addition, we introduce the asymmetry of the random error term using a Generalized Logistic (GL) distribution. Finally, in a different approach from that of Greene (1990), who assigns a Gamma function to the inefficiency component, we consider an Exponential distribution.

In some special cases, the convolution between the two error components admits a semi-closed expression also in cases of statistical dependence. An example is provided in Smith (2008), where the author obtains an expression for the density of the composite error in terms of hypergeometric functions.[1] We obtain a first generalization of Smith (2008) using a GL distribution for the random error density. This distribution describes situations of symmetry or asymmetry (positive or negative) according to values derived from one specific parameter. This allows us to analyze the statistical properties of a model which consider both statistical dependence and possible asymmetry in the random error component.

[1] Smith (2008) assigns an exponential distribution to the inefficiency error and a standard logistic distribution for the random error. It uses a Farlie-Gumbel-Morgerstern (FGM) copula to model the dependence between the two error components.

We are able to derive a semi-closed formula of the density function when we use a simple copula (FGM), but for more complex cases (i.e. when we use a more complicated copula such as the Frank copula), we build a computational tool that allows maximum likelihood estimation of SF models with a wide range of marginal distributions (see Nelsen (1999) for details on copula functions). The resulting approximations of the density of each sampling unit are then plugged into the log-likelihood function.

The paper proceeds as follows: Section 2 briefly explain the economic model and the statistical specification, and Section 3 reports the estimation from the US electricity industry, which were analyzed in previous research (Smith, 2008; Greene, 1990) to test for dependence in the case of Smith and to test for other marginal distributions in the case of Greene.[2] Conclusions follow.

2 Model specification

2.1 The economic model

The model to be estimated is a cost function, expressed by a Cobb-Douglas relationship,[3] with one output, Q, that is a function of three inputs, labor, capital and fuel, with respective factor prices, P_l, P_k and P_f.[4] In order to consider the homogeneity of cost function with respect to input prices, the dependent variable and two input prices (P_l and P_k) are expressed relative to P_f.

$$log\left(\frac{Cost}{P_f}\right) = \beta_1 logQ + \beta_2 log^2 Q + \beta_3 log\left(\frac{P_l}{P_f}\right) + \beta_4 log\left(\frac{P_k}{P_f}\right) + u + v, \quad (1)$$

2.2 The statistical model

In what follows, we report the proposition in which we derive the density function of the composite error ϵ when dependence is modeled through the FGM copula. For the more complicated Frank copula, we could not derive a closed formula of the function of the error term, but the numerical tool we implement allows us to obtain numerical estimations. We report details on the marginal distributions and copula functions in table I.

[2] The same data are used also by Christensen and Greene (1976), but we got data from Table 3 in Greene (1990).

[3] We estimate the same model as was employed by all the authors using the well-known dataset used in this paper.

[4] Also, a quadratic term for the output Q is introduced in the model.

In particular, we assume that $u \sim E(\delta_u)$, $v \sim GL(\alpha_v, \delta_v)$ and the dependence between u and v is modeled by FGM copula. Let $k_1(\epsilon)$ be defined as $k_1(\epsilon) = \exp\{-\frac{\epsilon+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}\}$, we derive the following:

- The density function of the composite error is

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

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

$$\begin{aligned} w_1(\epsilon) &= (1-\theta)\frac{\alpha_v k_1(\epsilon)^{-\alpha_v}}{\delta_u\big(\alpha_v+\frac{\delta_v}{\delta_u}\big)} & w_2(\epsilon) &= 2\theta\frac{\alpha_v k_1(\epsilon)^{-2\alpha_v}}{\delta_u\big(2\alpha_v+\frac{\delta_v}{\delta_u}\big)} \\ w_3(\epsilon) &= 2\theta\frac{\alpha_v k_1(\epsilon)^{-\alpha_v}}{\delta_u\big(\alpha_v+2\frac{\delta_v}{\delta_u}\big)} & w_4(\epsilon) &= -4\theta\frac{\alpha_v k_1(\epsilon)^{-2\alpha_v}}{\delta_u\big(2\alpha_v+2\frac{\delta_v}{\delta_u}\big)}. \end{aligned}$$

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

$$E[\epsilon] = \delta_u, \tag{3}$$

  and

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

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

Finally, the estimation of the cost efficiency CE_Θ is obtained through [5]

$$CE_\Theta = E[e^{-u}|\epsilon=\epsilon^*] = \frac{1}{f_\epsilon(.;\Theta)}\int_{\Re^+} e^{-u} f_{u,v}(u, x-u;\Theta)du. \tag{5}$$

[5] Details on calculation of CE scores are available upon request.

Table I: Marginal distribution functions and copulas.

	Parameters	Density	Distribution
Exponential	$\delta_u > 0$	$\frac{1}{\delta_u} e^{-\frac{u}{\delta_u}}$	$1 - e^{-\frac{u}{\delta_u}}$
GL	$\alpha_v,\ \delta_v > 0$	$\frac{\alpha_v}{\delta_v} \frac{e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}}{\left(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\right)^{\alpha_v+1}}$	$(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}})^{-\alpha_v}$
FGM copula	$\theta \in (-1, 1)$	$1 + \theta(1 - 2F_u)(1 - 2G_v)$	$F_u G_v\big(1 + \theta(1 - F_u)(1 - G_v)\big)$
Frank copula	$\theta \in (-\infty, \infty) \setminus \{0\}$	$\frac{\theta(1-e^{-\theta})e^{-\theta(F(u)+G(v))}}{[(1-e^{-\theta})-(1-e^{-\theta F(u)})(1-e^{-\theta G(v)})]^2}$	$-\theta^{-1} \ln[1 + \frac{(e^{-\theta F(u)}-1)(e^{-\theta G(v)}-1)}{(e^{-\theta}-1)}]$

3 Empirical results

Our empirical application concerns the estimation of cost frontier for a sample of 123 firms operating in US electricity markets in 1970. As mentioned in the Introduction, the same data sample was used by Greene (1990) and Smith (2008) to provide applications of new SF specifications. We include these applications in order to provide a thorough comparison between the different statistical models. In fact, in addition to the classic SF, we estimate different models, as summarized in Table II.

Table II: Summary of the statistical models.

Name	Random Error Distribution	Inefficiency Distribution	Dependence
Classic SF	Normal (σ_v^2)	Truncated Normal (σ_u^2)	No
IS	Symmetric GL ($\alpha_v = 1, \delta_v$)	Exp (δ_u)	No
DS	Symmetric GL ($\alpha_v = 1, \delta_v$)	Exp (δ_u)	FGM copula
DS$_{Frank}$	Symmetric GL ($\alpha_v = 1, \delta_v$)	Exp (δ_u)	Frank copula
IA	GL (α_v, δ_v)	Exp (δ_u)	No
DA	GL (α_v, δ_v)	Exp (δ_u)	FGM copula
DA$_{Frank}$	GL (α_v, δ_v)	Exp (δ_u)	Frank copula

Legend: $ClassicSF$, the traditional model of SF; IS, independence and symmetry; DS, model with FGM dependence and symmetry; DS_{Frank}, model with Frank dependence and symmetry; IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; DA_{Frank}, model with Frank dependence and asymmetry.

We use classic SF as the benchmark model. All the other models have the same marginal distributions of the inefficiency error, while the distribution of random error component changes depending on which specification is considered. In detail, IS, DS and DS$_{Frank}$ models consider the symmetry of the random error term (the parameter α_v, which is the skewness measure of GL distribution, is equal to 1, meaning symmetry); IA, DA and DA$_{Frank}$ models consider the possibility of variation in the sign of skewness of the random error.

Within these two groups of models, specifications differ from each other based on the functional form of dependence and the skewness of v. In particular, we estimate one model with no dependence, one specification with the FGM copula and one model with the Frank copula.

We report the empirical results in table III. The t-statistics are shown below the estimated coefficients.

After obtaining significant elasticities, we focus on the measures of association θs, i.e. the copula parameters, that are negative but not significant in all three models constructed under hypotheses of dependence. This is not a surprising finding because Smith (2008) also rejects the dependence between u and v. Turning to the skewness issue, which is measured through α_v-parameters, table IV shows the t-test on symmetry of the random error term v. In particular, we widely accept the null hypothesis of symmetry in all three models (IA, DA and DA$_{Frank}$) could capture the possible asymmetry of the random error term. These two results are in line with the choice of IS model. In fact, it shows the smallest value of AIC measure, even if, following Burnham and Anderson (2004), IS, DS and IA models are indifferent each other.

Based on the estimated models we compute the individual cost efficiencies (CE). In table V, we report descriptive statistics of CE for each model (i.e. mean, standard deviation, minimum and maximum values), and figure 1 shows the plot of CE for each firm and for the different specification. Figure 2 illustrates the empirical density functions of efficiency levels for every model estimation. From both table and figures, we can see the heterogeneity of results among the different models. For the IS model, which produces the best specification, we can see that the estimated levels of CE are higher than those calculated through classic SF. The average for the latter is 0.89, while is 0.94 for the former. The IS estimations show more variability than Classic SF. In addition, the estimated values tend to be similar in the same class of models, especially when dependence is the characteristic used to identify the class. In summary, figures 1 and 2, and table V show that the efficiency scores estimated using our specifications are quite different among different models, signaling the presence of an effect of dependence between u and v and/or asymmetry of random error.

Table III: Estimates for US electricity market.

	Classic SF	IS	DS	DS$_{Frank}$	IA	DA	DA$_{Frank}$
β_0	**-7.410**	**-7.877**	**-7.800**	**-7.875**	**-7.786**	**-7.773**	**-7.789**
	-22.17	*-25.45*	*-25.16*	*-26.68*	*-23.01*	*-24.36*	*-26.29*
β_1	**0.408**	**0.4467**	**0.4472**	**0.4712**	**0.4489**	**0.4502**	**0.4633**
	10.32	*12.84*	*13.01*	*16.00*	*13.05*	*13.04*	*15.83*
β_2	**0.031**	**0.0283**	**0.0283**	**0.0269**	**0.0282**	**0.0281**	**0.0274**
	11.55	*11.71*	*11.86*	*13.06*	*11.66*	*11.63*	*13.25*
β_3	**0.245**	**0.3111**	**0.2904**	**0.2855**	**0.2953**	**0.2870**	**0.2747**
	3.70	*4.97*	*4.63*	*4.82*	*4.27*	*4.43*	*4.60*
β_4	0.059	0.0236	0.0329	0.0221	0.0356	0.0360	0.0268
	0.96	*0.42*	*0.59*	*0.41*	*0.62*	*0.64*	*0.50*
δ_u		**0.097**	**0.123**	**0.133**	**0.107**	**0.129**	**0.133**
		4.15	*3.02*	*10.05*	*4.60*	*1.69*	*10.04*
α_v					**0.662**	**0.745**	**1.020**
					1.50	*1.73*	*2.61*
δ_v		**0.058**	**0.064**	**0.069**	**0.045**	**0.055**	**0.069**
		6.28	*4.56*	*11.52*	*2.20*	*2.36*	*6.37*
θ			-0.99984	-0.4973		-0.99980	-0.4083
			-0.85	*-0.19*		*-0.41*	*-0.17*
$\gamma = \frac{V(u)}{V(\epsilon)}^*$	0.673	0.462	0.726	0.790	0.542	0.771	0.801
log-likelihood	66.12	68.20	68.62	67.54	68.52	68.74	67.51
AIC	-118.24	-122.39	-121.25	-119.08	-121.04	-119.47	-117.02

Source: our elaborations on data from Greene (1990).

Legend: $ClassicSF$, the traditional model of SF; IS, independence and symmetry; DS, model with FGM dependence and symmetry; DS_{Frank}, model with Frank dependence and symmetry; IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; DA_{Frank}, model with Frank dependence and asymmetry.

*V(ϵ) for DS_{Frank} and DA_{Frank} is calculated as the variance of the estimated $\hat{\epsilon}$.

Table IV: Results of t-test on symmetry for v.

$H_0 : \alpha_v = 1$ vs $H_1 : \alpha_v \neq 1$

	IA	DA	DA$_{Frank}$
t-statistic	-0.766	-0.592	0.051
p-value	0.444	0.552	0.999

Source: our elaborations on data from Greene (1990).

Table V: Some descriptive statistics of cost efficiency.

	Classic SF	IS	DS	DS$_{Frank}$	IA	DA	DA$_{Frank}$
MEAN	0.8884	0.9447	0.8956	0.8815	0.9413	0.8913	0.8792
SD	0.0536	0.0627	0.0657	0.0605	0.0735	0.0479	0.0668
MIN	0.6812	0.6286	0.6271	0.5674	0.5891	0.6195	0.5693
MAX	0.9704	0.999957	0.9980	0.9519	0.9999	0.9973	0.9530

Source: our elaborations on data from Greene (1990).

Legend: $ClassicSF$, the traditional model of SF; IS, independence and symmetry; DS, model with FGM dependence and symmetry; DS_{Frank}, model with Frank dependence and symmetry; IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; DA_{Frank}, model with Frank dependence and asymmetry.

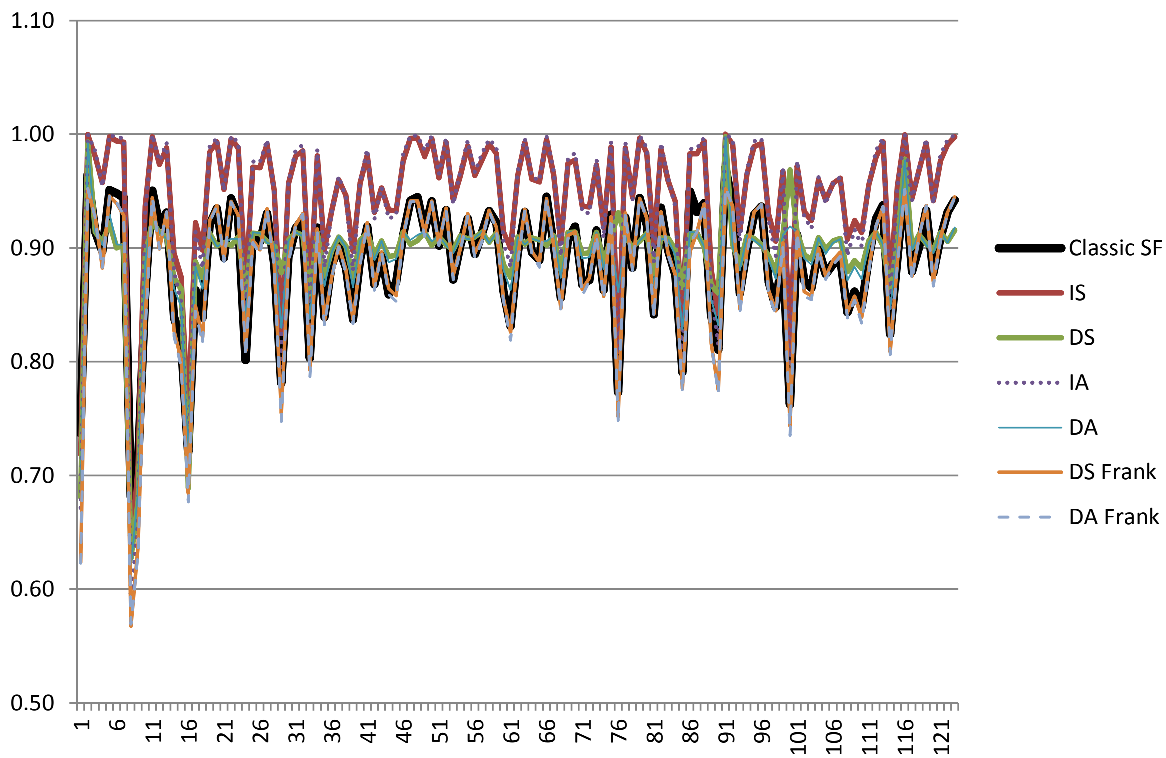
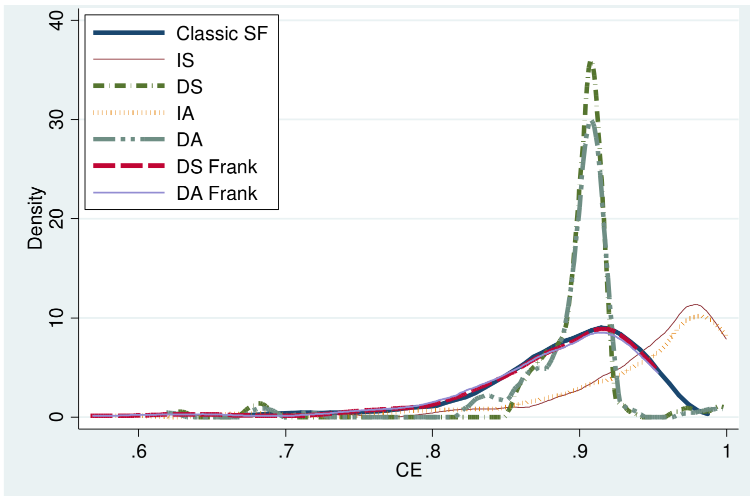


Figure 1: Cost efficiency by observation for each model.

Figure 2: Kernel density of cost efficiency for each model.



4 Conclusions and future research

In this paper we propose the derivation of cost SF following the approach of Bonanno et al. (2017), which models dependence between the random error and inefficiency components through copula functions and assigns a flexible distribution to random error. The best model for the sample of 123 firms operating in the US electricity market in 1970 is the most simple IS specification (independence and symmetry), but in terms of CE scores, we estimate different values than those obtained through classic SF. Evidences proves that our models offer improvement over classic SF estimations.

Our findings suggest additional research to further develop SF models would be beneficial. In particular, one aspect to consider is the assignment of a more flexible distribution for the inefficiency component. When exponential function is used for u, the so-called γ-parameter, calculated as the ratio between the variance of the inefficiency and the variance of the overall error term, depends on the estimated value of δ_u, to a significant extent, as the variance of *u-exponential* is equal to δ_u^2. In our empirical application on data from US electricity, we estimate a "good value" of δ_u, which allows us to obtain a γ-parameter signaling the presence of inefficiency, but, in many cases, the value of δ_u could be very small. However, the conclusion of the absence of inefficiency, in this case, could be misleading.

References

Bonanno, G., D. De Giovanni, and F. Domma (2017). The "wrong skewness" problem: A re-specification of stochastic frontiers. *Journal of Productivity Analysis Vol. 47 (1)*, 49–64.

Burnham, K. and D. Anderson (2004). Multimodel Inference. understanding AIC and BIC in Model Selection. *Sociological Methods Research 32*(2), 261–304.

Christensen, L. and W. Greene (1976). Economies of Scale in U.S. Electric Power Generation. *The Journal of Political Economy 84*(4), 655–676.

Greene, W. H. (1990). A Gamma-distributed stochastic frontier model. *Journal of Econometrics 46*(1–2), 141–163.

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

Nelsen, R. B. (1999). *An Introduction to Copula*. New York: Springer.

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