Cost Efficiency, Asymmetry and Dependence in US electricity industry.

Graziella Bonanno

bonanno@diag.uniroma1.it

Department of Computer, Control, and Management Engineering ”Antonio Ruberti” - Sapienza University of Rome (Italy)

Cost Efficiency, Asymmetry and Dependence in US electricity industry.

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.

Key words:Stochastic frontier, cost, Copula functions, skewness, dependence.

1 Introduction

After the methodological derivation of a new specification of SF proposed for production functions by Bonanno et al. (2017), this paper aims to show how asymmetry of the random error and dependence between it and the inefficiency component are introduced also in cost frontiers.

The basic formulation of a cost frontier model can be expressed as $c = f(\mathbf{Q}, \mathbf{P}; \boldsymbol{\beta})e^{\epsilon}$, where c are the firm-specific total costs, $\mathbf{Q}$ is a vector of outputs, $\mathbf{P}$ is a vector of input prices and $\boldsymbol{\beta}$ is the vector of unknowns parameters (details are in Kumbhakar and Lovell, 2000). The error term, ϵ, 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 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 $\epsilon = u + v$ in a cost function. Moreover, it allows random variation of frontier across firms.

We propose a model can capture the dependence structure between u and v, modeling it with a copula function that allows 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 assigning it a Generalized Logistic distriution. Finally, differing from 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 for the model with an exponential distribution for the inefficiency error, a logistic distribution for the random error and a FGM copula. We obtain a first generalization of Smith (2008) by using a Generalized Logistic (GL) distribution for the random error density. This distribution describe situations of symmetry or asymmetry (positive or negative) according to values that takes on one of its parameters. This allows us to analyze the statistical properties of a model in which both statistical dependence and possible asymmetry in the random error component.

We can derive the explicit density function when we use a simple copula (FGM), but for more complex cases (i.e. when we use a more complicated copula, the so-called 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 show briefly the economic model and the statistical specification and the Section 3 reports the estimation from the US electricity industry already analyzed in earlier paper by Smith (2008) to test for dependence and in paper of Greene (1990) to test other marginal distributions.[1] Conclusions follow.

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

2 The model specification

2.1 The economic model

The model to be fit is a cost function, expressed by a Cobb-Douglas relationship, with one output Q that is a function of three labor, capital and fuel, with respective factor prices P_l, P_k and P_f. 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. The statistical model we use, however, includes a location parameter that can lead to identification problems. For this reason, differently from Smith, we do not include the intercept β_0. Smith (2008) use a normal–half normal model for the marginal distributions, while our choice is based on a more flexible distribution of random error (Table 2).

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

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 FGM copula. For the more complicated Frank copula, we cannot derive an explicit function of the error term, but we use a numerical tool in order to obtain the estimations. We report details on the marginal distributions and copula functions in table 1.

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:

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

- 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.

Table 1: 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	$\alpha_v,\ \delta_v > 0$	$\frac{\alpha_v}{\delta_v}\frac{e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}}{\big(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}}\big)^{\alpha_v+1}}$	$(1+e^{-\frac{v+\delta_v[\Psi(\alpha_v)-\Psi(1)]}{\delta_v}})^{-\alpha_v}$
FGM copula	$\theta \in (-1,1)$	$1+\theta(1-2F_u)(1-2G_v)$	$F_uG_v\big(1+\theta(1-F_u)(1-G_v)\big)$
Frank copula	$\theta \in (-\infty,\infty)\setminus\{0\}$	$\frac{\theta(1-e^{-\theta})e^{-\theta(F(u)+G(v))}}{[(1-e^{-\theta})-(1-e^{-\theta F(u)})(1-e^{-\theta G(v)})]^2}$	$-\theta^{-1}\ln[1+\frac{(e^{-\theta F(u)}-1)(e^{-\theta G(v)}-1)}{(e^{-\theta}-1)}]$

Finally, the estimation of the cost efficiency CE_Θ is obtained through [2]

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

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

3 Empirical results

Our application concerns the estimation of cost frontier for a sample of 123 firms operating in US electricity markets in 1970. As we already mentioned in the Introduction, the same data sample was used by Greene (1990) and Smith (2008) to give applications of new SF specifications. We include this application in order to provide a thorough comparison between different statistical models. In fact, in addition to the classic SF, we estimate different models as summarised in Table 2.

Table 2: Summary of the statistical models.

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

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

Classic SF is used as benchmarking model. All the other models have the same marginal distributions of the inefficiency error, while the distribution of random error component changes depending on whether we consider v-symmetric or v-asymmetric.

In detail, models differ each other by the functional form of dependence and the skewness of v. In particular, we fit one model with no dependence, one with the FGM copula and one model with the Frank copula. Then, considering $\alpha_v = 1$, we estimate the same three models as above (with independence, FGM and Frank copulas).

Table 3 reports the results. In parentheses, there are the t-statistics.

After obtaining significative elasticities, the attention is for the measure of association θ that is negative but not significative in all three models constructed under hypotheses of dependence. It is not a surprising finding because also Smith (2008) rejects the dependence between u and v.

Now the attention is for α_v-parameter. Table 4 shows the t-test on symmetry of v. In particular, we widely accept the null hypothesis of symmetry in all three models can capture the possible asymmetry of the random error (IA, DA and DA$_{Frank}$).
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 efficiency (CE). In table 5 we report some descriptive statistics of CE for each mode (i.e. mean and standard deviation)l and figure 1 shows plot of the estimated efficiencies of each firm for all models we fit. From both, we can see eterogeneity of results across different models. 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.

Table 3: 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: IS stands for independence and symmetry; DS is the model with FGM dependence and symmetry; DS_{Frank} stands for Frank dependence and symmetry; IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; DA_{Frank} stands for Frank dependence and asymmetry.
*V(ϵ) for DS_{Frank} and DA_{Frank} is calculated as the variance of the estimated $\hat{\epsilon}$.

Figure 1: Cost Efficiency by observation for each model

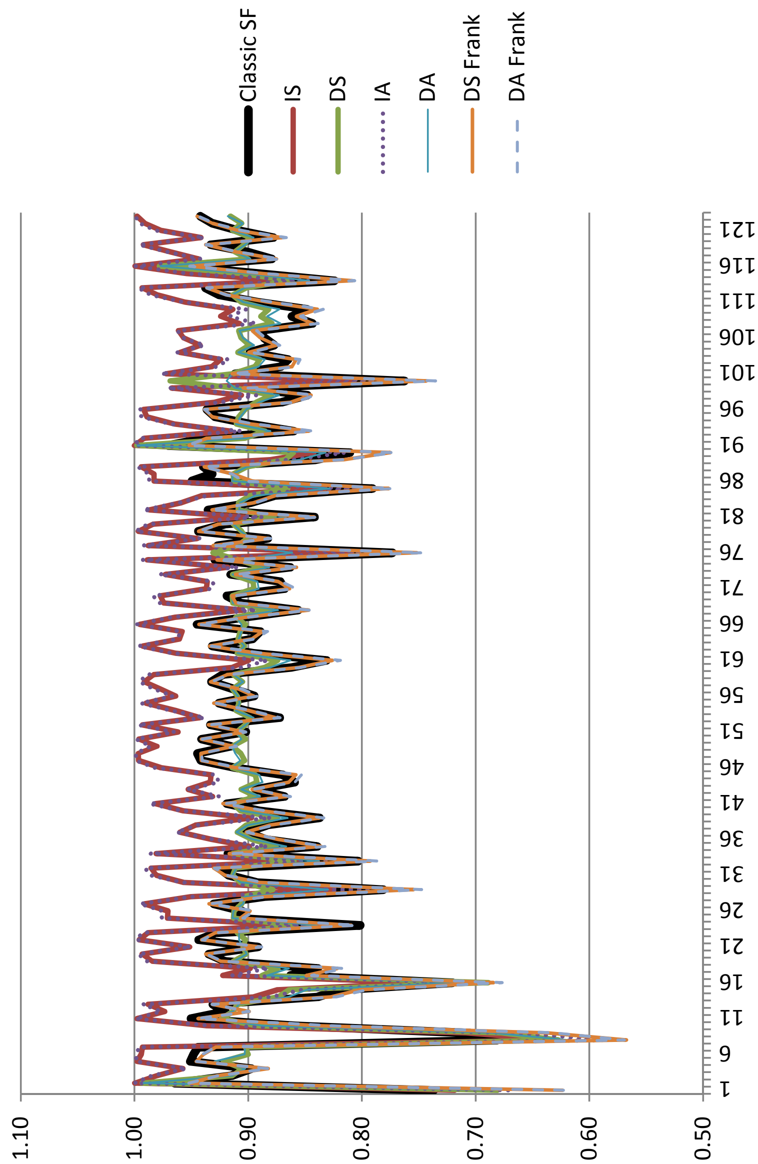

Figure 2: Kernel density of Cost Efficiency for each model

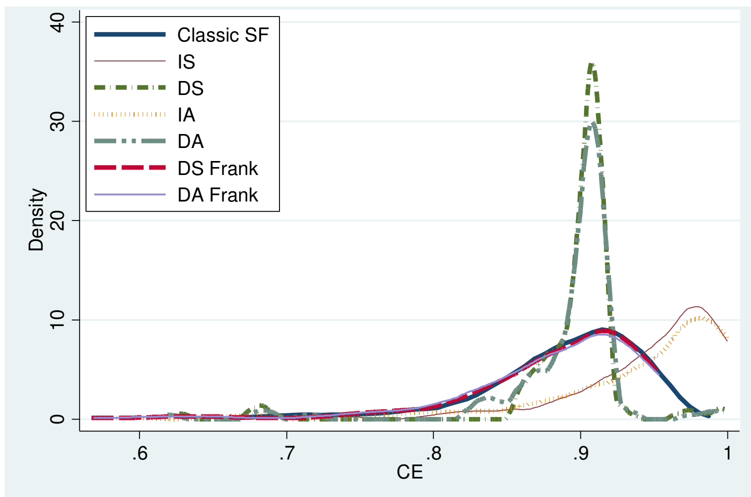

Table 4: Results of t-test on symmetry for v.

$H_0 : \alpha_v = 1$ vs $H_1 : \alpha_v \neq 1$			
	IA	DA	DA$_{Frank}$
t-statistic	0.796	0.514	0.051
p-value	0.401	0.660	0.9601

Table 5: 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: IS stands for independence and symmetry; DS is the model with FGM dependence and symmetry; DS_{Frank} stands for Frank dependence and symmetry; IA is the model with independence and asymmetry; DA stands for FGM dependence and asymmetry; DA_{Frank} stands for Frank dependence and asymmetry.

4 Conclusions and future research

Even if we reject dependence and asymmetry, our models improve the classic SF estimations. This finding suggests to continue researching for further evolution of SF models. In particular, one aspect to consider is the assignment of a more flexible distribution also for the inefficiency component. When Exponential 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 in significative 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, δ_u tend to be very small. The conclusion of the absence of inefficiency in this case could be misleading. This is the reason to propose a more general model in which also u is distributed through a more flexible function.

References

G. Bonanno, D. De Giovanni, and F. Domma. The wrong skewness problem: A re-specification of stochastic frontiers. *Journal of Productivity Analysis*, forthcoming, 2017.

K.P. Burnham and D.R. Anderson. Multimodel Inference. understanding AIC and BIC in Model Selection. *Sociological Methods Research*, 32(2):261304, 2004.

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

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

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

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

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