J Prod Anal (2013) 40:1–14
DOI 10.1007/s11123-012-0289-8

Maximum likelihood estimation of seemingly unrelated stochastic frontier regressions

Hung-pin Lai · Cliff J. Huang

Published online: 16 June 2012
© Springer Science+Business Media, LLC 2012

Abstract In this paper, we propose the copula-based maximum likelihood (ML) approach to estimate the multiple stochastic frontier (SF) models with correlated composite errors. The motivation behind the extension to system of SF regressions is analogous to the classical generalization to system of seemingly unrelated regressions (Zellner in J Am Statist Assoc 57:348–368, 1962). A demonstration of the copula approach is provided via the analysis of a system of two SF regressions. The consequences of ignoring the correlation between the composite errors are examined by a Monte Carlo experiment. Our findings suggest that the stronger the correlation between the two SF regressions, the more estimation efficiency is lost in separate estimations. Estimation without considering the correlated composite errors may cause significantly efficiency loss in terms of mean squared errors in estimation of the SF technical efficiency. Finally, we also conduct an empirical study based on Taiwan hotel industry data, focusing on the SF regressions for the accommodation and restaurant divisions. Our results, which are consistent with the findings in simulation, show that joint estimation is significantly different from separate estimation without considering the correlated composite errors in the two divisions.

Keywords Maximum likelihood estimation · Copula · Seemingly unrelated stochastic frontier regressions

H. Lai (✉)
Department of Economics, National Chung Cheng University, Chia-Yi, Taiwan
e-mail: ecdhpl@ccu.edu.tw

C. J. Huang
Department of Economics, Vanderbilt University, Nashville, TN, USA

JEL Classification C3 · C5 · R3

1 Introduction

Since the pioneering work of Aigner et al. (1977), stochastic frontier (SF) analysis has been widely used in productivity and efficiency studies to describe and estimate production, cost, and profit frontier models. Such analysis typically assumes that a decision-making unit (DMU) employs a single production process or technology in a production of an output using multi-inputs. Most of the econometric techniques have focused on the estimation and interpretation of a single stochastic frontier regression with the composite error on inefficiency and statistical noise. There has been, however, very little discussion in the stochastic frontier literature to extend the single SF regression to multiple SF regressions settings, which would allow for the possibility of correlation among composite errors across SF regressions. The motivation behind the extension to system of SF regressions is analogous to the classical generalization to system of seemingly unrelated regressions (Zellner 1962).

Organizations or DMUs often operate multiple production divisions or sub-DMUs, with each division supported by its own set of resource inputs. A tourist hotel, for example, may provide services of accommodation and restaurants under separate divisions within the same DMU hotel management. Given that these divisions would still belong to the same DMU, they may be subject to the same random shocks as the parent DMU; and given that these divisions would share some commonly observed or unobserved characteristics of the parent DMU, the divisions' technical efficiencies may well be correlated. In this circumstance, a system of multiple stochastic frontier regressions on the sub-DMUs is a more appropriate representation of a DMU's operation and performance.

In this paper, the stochastic frontier model under consideration focuses on the correlation among a set of individual SF regressions, and still assumes independence between the error components: the statistical noise and the inefficiency component. The correlation may come from the correlated statistical noises or the correlated inefficiency components, or both. This pattern would therefore seem to suggest that the estimation of these stochastic frontier regressions should take into account the level of mutual dependency among them, as opposed to the regression-by-regression estimation.

Lately, several studies have used the copula method in studying of stochastic frontier models. Amsler et al. (2011) use copulas to model time dependence in stochastic frontier models. Shi and Zhang (2011) consider a copula regression model for dependence in long-tail distributions. Carta and Steel (2010) suggest using copulas in modeling multi-output stochastic frontiers. Various types of dependence of SF models are investigated with utilizing the copula method. Smith (2008), on the other hand, proposes a copula function to link the marginal distributions of correlated components v and u in order to construct the distribution of the composite error ε.

Copulas, originally introduced by Sklar (1959), provide a useful method of deriving joint distribution of a set of random variables given the univariate margins, especially in the case of nonnormal margins of the SF composite errors. However, the copula-based likelihood function of the joint SF regressions requires the computation of the cumulative distribution function (CDF) of the composite error. Such computation typically requires numerical integration procedures, which proves to be onerous. To facilitate the computation of the copula-based maximum likelihood estimation, we follow Tsay et al. (2012) in deriving an easy to implement and accurate closed-form formula for computing the CDF of the composite error.

A demonstration of our approach is provided through the analysis of a set of two SF regressions. The consequences of ignoring the dependence between the two SF regressions are examined via a Monte Carlo simulation, where both efficiency and bias in estimation are investigated under various degrees of correlation. The findings of the simulated experiment suggest that the efficiency gain in estimation from the system of SF regressions over the regression-by-regression increases significantly as the correlated composite errors become stronger. The efficiency gain is particularly significant in estimating the variances of the symmetric and one-sided errors, which is crucial in predicting the measures of the technical efficiency.

Finally, a two-SF regressions model is applied to an empirical study of the hotel industry in Taiwan, focusing on the production efficiency in accommodation and restaurant divisions. Our results show that both the estimated production frontiers and technical efficiencies differ significantly between the system estimation and the regression-by-regression estimations. They suggest that ignoring the interdependence in operations of the two divisions may cause a downward biased estimator of the technical efficiency.

While we apply a set of SF regressions to the divisions or sub-DMUs within a DMU, the accommodation and restaurant divisions of 50 international grand hotels, we recognize that the procedure is more generally applicable. For example, it can be applied to the SF analysis of temporal cross-section data where the "sub-DMUs" are the departments of a bank, or an efficiency study of the airline industry where the "sub-DMUs" are firms with correlated composite errors. The application may extend to macroeconomic settings where the "sub-DMUs" are countries within the same economic region or block.

The plan of the paper is as follows. In Sect. 2 we describe a multiple stochastic frontier regressions model with correlated composite errors. Section 3 proposes a copula-based joint maximum likelihood estimation of the frontier regressions. We then turn to the computation issue of evaluating the Gaussian copula-based likelihood function in Sect. 4. The comparison of the proposed joint multiple regressions estimation versus the separated regression-by-regression is studied via a Monte Carlo simulation in Sect. 5. An empirical study of the hotel industry in Taiwan is given in Sect. 6. Lastly, a summary conclusion is given in Sect. 7.

2 A multiple stochastic frontier regressions model

Suppose the production operations of a DMU consist of J sub-DMUs (divisions), each of which produces a single output under the Cobb-Douglas type of production technology. The production frontiers of the J sub-DMUs are represented by a set of SF regressions,

$$\begin{aligned} y_{1i} &= x_{1i}^{\mathrm{T}}\beta_1 + v_{1i} - u_{1i}, \\ y_{2i} &= x_{2i}^{\mathrm{T}}\beta_2 + v_{2i} - u_{2i}, \quad i = 1, 2, \ldots, N \\ &\vdots \\ y_{Ji} &= x_{Ji}^{\mathrm{T}}\beta_J + v_{Ji} - u_{Ji}, \end{aligned} \tag{1}$$

where y_{ji} and x_{ji} respectively denote the log output and the log inputs of the jth sub-DMU ($j = 1, 2, \ldots, J$) of the DMU i; $v_{ji} \sim N(0, \sigma_{v_j}^2)$ and $u_{ji} \sim N^{+}(0, \sigma_{u_j}^2)$ respectively represent the noise component and the non-negative inefficiency component. For a given j, the sub-DMU's v_{ji} and u_{ji} are assumed to be mutually independent. Thus, for any given SF regression, the noise component is uncorrelated with the inefficiency component, which is a standard assumption in a single stochastic frontier modeling.[1]

When the composite error is defined as $\varepsilon_{ji} = v_{ji} - u_{ji}$, the correlations among the composite errors ε_{ji} are the consequence of the correlation in v_{ji} and the correlation in u_{ji}. Across the SF regressions in (1), u_{ji} and u_{si} are correlated due to the sharing of the same common characteristics among the J sub-DMUs within a DMU. Similarly, across the SF regressions, v_{ji} and v_{si} are allowed to be correlated for $j \neq s$, possibly due to common stochastic shocks to the DMU and its sub-DMUs. Conversely, however, in our model, the correlation of the composite errors is not decomposable to the correlation in v_{ji} or to the correlation in u_{ji}. Nevertheless, we should be able to show that the SF regression-by-regression estimation of (1) is less efficient than the system estimation, especially for $\sigma^2_{v_j}$ and $\sigma^2_{u_j}$.

In estimating a standard single sector's stochastic frontier (SF) regression model using the maximum likelihood (ML) approach, it is often assumed that the noise component v_{ji} follows a normal distribution, and that the inefficiency component u_{ji} follows a half-normal (or truncated-normal) distribution. In this case, the derivations of the probability density function (PDF) of the composite error ε_{ji} and the ML estimator of a single SF regression are relatively straightforward. However, computation of the likelihood function in the joint estimation of (1) with correlated composite errors is sometimes tedious in the analytic form if no further simplified assumptions are imposed. For instance, Wang and Ho (2010) assume the scaling property in the distribution of u to simplify the model. Pitt and Lee (1981) use a multivariate truncated normal distribution but the computation of likelihood function involves T-dimensional integrals, which requires numerical approximation and the precision of the approximation may be an issue. We therefore propose the copula approach as the means of constructing the joint probability of $\varepsilon_{1i}, \ldots, \varepsilon_{Ji}$ in (1). In Sect. 3 we begin with some basics on the copula, and then construct the likelihood function of the system of SF model by the copula approach.

[1] This independence assumption of the error components in a single SF regression setting has been relaxed in the studies by Pal and Sengupta (1999), Pal (2004), Bandyopadhyay and Das (2006), and Smith (2008). However, these studies are all confined to a single regression case.

3 Copula and the ML estimator

Let $\theta_j = (\beta_j^{\mathrm{T}}, \sigma_{vj}, \sigma_{uj})^{\mathrm{T}}$ be a vector of parameters in the jth SF regression, and $F_j(\varepsilon_{ji}) = F(\varepsilon_{ji}; \theta_j)$ be the cumulative distribution function (CDF) of the composite error of the jth SF regression with the vector of parameters θ_j. In order to simplify the notation in our following analysis, we suppress θ_j in the marginal CDF $F(\varepsilon_{ji}; \theta_j)$ and use the subscript j to indicate the jth SF regression, so $F_j(\varepsilon_{ji})$ will be used instead in our following analysis.

In deriving the joint CDF of J random variables, we rely on the results of Sklar's theorem (1959), and Schweizer and Sklar (1983), in which the joint CDF can be expressed as a function of its own one-dimensional margins. The function is the copula. More specifically, the joint CDF of the composite errors $\varepsilon_i = (\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$ from the J SF regressions can be represented by a J-copula function $C(\cdot)$,

$$F(\varepsilon_{1i}, \ldots, \varepsilon_{Ji}) = C(F_1(\varepsilon_{1i}), \ldots, F_J(\varepsilon_{Ji}); \rho) \tag{2}$$

where ρ is a vector of parameters of the copula called the dependence parameter, which measures dependence between the marginal CDFs. Thus, the joint distribution of the composite errors $\varepsilon_i = (\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$ is expressed in terms of the marginal distributions, and the copula function binds them together. The advantage of copula function representation in (2) is that, while each SF regression in (1) corresponds to a separate $F_j(\varepsilon_{ji})$, it allows us to identify the dependence among the multiple SF regressions through the dependence in their marginal CDFs. Therefore, the copula approach is an appropriate technique of studying the correlated stochastic frontier regressions. It has been shown that the copula function $C(\cdot)$ is unique if $F_1(\cdot), \ldots, F_J(\cdot)$ are all continuous.[2] Furthermore, since $0 \le F_j(\varepsilon_{ji}) \le 1$, the copula function can be viewed as multivariate distribution of uniform $U[0, 1]$ variables with the dependence parameter ρ.

In general, different copula functions capture different dependence structures. One special case is the product copula where $C(F_1(\varepsilon_{1i}), \ldots, F_J(\varepsilon_{Ji}); \rho) = \prod_{j=1}^{J} F_j(\varepsilon_{ji})$. It follows then, from (2), that $F(\varepsilon_{1i}, \ldots, \varepsilon_{Ji}) = \prod_{j=1}^{J} F_j(\varepsilon_{ji})$. Hence, the product copula implies the mutual independence among ε_{1i}, …, ε_{Ji}, or the independence among the J SF regressions.

For the maximum likelihood estimation, the joint PDF, $f(\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$, instead of the joint CDF, $F(\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$, is required. By taking the derivatives of (2) with respect to ε_{1i}, …, ε_{Ji}, the corresponding joint PDF is obtained,

[2] See section 2.2.2 of Trivedi and Zimmer (2005).

$$f(\varepsilon_{1i},\ldots,\varepsilon_{Ji}) = c(F_1(\varepsilon_{1i}),\ldots,F_J(\varepsilon_{Ji});\rho) \times \prod_{j=1}^{J} f_j(\varepsilon_{ji}), \quad (3)$$

where $c(F_1(\varepsilon_{1i}),\ldots,F_J(\varepsilon_{Ji});\rho) = \frac{\partial^J C(F_1(\varepsilon_{1i}),\ldots,F_J(\varepsilon_{Ji});\rho)}{\partial F_1(\varepsilon_{1i}),\ldots,\partial F_J(\varepsilon_{Ji})}$ is the copula density and $f_j(\varepsilon_{ji})$ is the marginal PDF. In the case of product copula, we have $c(\cdot) = 1$, and it follows from (3) that the joint PDF is the product of the marginal PDFs.

Several multivariate copulas can be used in (2) under this framework. Further extensions of the proposed approach to other copula functions should be able to follow the same procedure with a similar calculation. For instance, the multivariate Student's t copula, Archimedean copula, Gumble n-coplua, and Clayton n-copula, see Cherubini et al. (2004) for a systematic review of the copula functions.

For the joint estimation of the multiple SF regressions in (1), we use the Gaussian copula to derive the J-dimensional distribution function $F(\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$ in (2). The Gaussian copula takes the form[3]

$$\begin{aligned} C(\gamma_{1i},\ldots,\gamma_{Ji};\Omega) &= \Phi_J\left(\Phi^{-1}(\gamma_{1i}),\ldots,\Phi^{-1}(\gamma_{Ji});\Omega\right), \\ &= \int_{-\infty}^{\Phi^{-1}(\gamma_{1i})} \cdots \int_{-\infty}^{\Phi^{-1}(\gamma_{Ji})} \\ &\quad \times \frac{1}{(2\pi)^{J/2}|\Omega|^{1/2}} e^{-\frac{1}{2}z^{\mathrm{T}}(\Omega^{-1})z} dz_1\ldots dz_J, \end{aligned} \quad (4)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution, and $\Phi_J(\cdot)$ is the CDF of a standard J-variate normal distribution of the random variables with the $J \times J$ correlation matrix $\Omega = [\Omega_{js}]$, i.e., the diagonal elements are $\Omega_{jj} = 1$, and the off-diagonal elements Ω_{js} are the correlation coefficients between two variables, $\Phi^{-1}(\gamma_{ji})$ and $\Phi^{-1}(\gamma_{si})$. The corresponding Gaussian copula density of (4) is

$$c(\gamma_{1i},\ldots,\gamma_{Ji};\Omega) = \frac{1}{|\Omega|^{1/2}} e^{-\frac{1}{2}\zeta_i^{\mathrm{T}}\left(\Omega^{-1}-I_J\right)\zeta_i}, \quad (5)$$

where $\zeta_i = \left(\Phi^{-1}(\gamma_{1i}),\ldots,\Phi^{-1}(\gamma_{Ji})\right)^{\mathrm{T}}$ and I_J is a $J \times J$ identity matrix.

Replacing $\gamma_{ji} = F_j(\varepsilon_{ji})$ in (4), the joint CDF of the composite errors in (2) becomes,

$$F(\varepsilon_{1i},\ldots,\varepsilon_{Ji}) = \Phi_J\left(\Phi^{-1}(F_1(\varepsilon_{1i})),\ldots,\Phi^{-1}(F_J(\varepsilon_{Ji}));\Omega\right). \quad (6)$$

The corresponding joint PDF of the composite errors in (3) becomes,

$$\begin{aligned} f(\varepsilon_{1i},\ldots,\varepsilon_{Ji}) &= c(F_1(\varepsilon_{1i}),\ldots,F_J(\varepsilon_{Ji});\Omega) \times \prod_{j=1}^{J} f_j(\varepsilon_{ji}) \\ &= \frac{1}{|\Omega|^{1/2}} e^{-\frac{1}{2}\zeta_i^{\mathrm{T}}\left(\Omega^{-1}-I_J\right)\zeta_i} \times \prod_{j=1}^{J} f_j(\varepsilon_{ji}) \end{aligned} \quad (7)$$

where $\zeta_i = \left(\Phi^{-1}(F_1(\varepsilon_{1i})),\ldots,\Phi^{-1}(F_J(\varepsilon_{Ji}))\right)^{\mathrm{T}}$. Note that the off-diagonal elements of Ω measure the correlation coefficients between two variables, $\Phi^{-1}(F_j(\varepsilon_{ji}))$ and $\Phi^{-1}(F_s(\varepsilon_{si}))$. Thus, if the off-diagonal elements of Ω are all zeros, Ω becomes a $J \times J$ identity matrix, i.e., $\Omega = I_J$. In this case, the Gaussian copula density $c(\cdot) = 1$ and $f(\varepsilon_{1i},\ldots,\varepsilon_{Ji}) = \prod_{j=1}^{J} f_j(\varepsilon_{ji})$, which implies the mutual independence of the composite errors $\varepsilon_i = (\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$, and hence the mutual independence of the SF regressions in (1).

Here, the dependence parameter ρ in (2) is defined as the column vector obtained by stacking the column vectors of the off-diagonal lower triangular matrix of Ω in (4). Therefore, based on (7) we may write the log-likelihood function of the J multiple SF regressions of (1) as

$$\begin{aligned} \ln L(\theta) &= \sum_{i=1}^{N} \ln f(\varepsilon_{1i},\ldots,\varepsilon_{Ji}), \\ &= \sum_{i=1}^{N} \ln c(F_1(\varepsilon_{1i}),\ldots,F_J(\varepsilon_{Ji});\Omega) + \sum_{i=1}^{N}\sum_{j=1}^{J} \ln f_j(\varepsilon_{ji}) \end{aligned} \quad (8a)$$

$$\begin{aligned} &= -\frac{N}{2}\ln|\Omega| - \frac{1}{2}\sum_{i=1}^{N} \zeta_i^{\mathrm{T}}\left(\Omega^{-1} - I_J\right)\zeta_i \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{J} \ln f_j(\varepsilon_{ji}),, \end{aligned} \quad (8b)$$

where $\theta = \left(\theta_1^{\mathrm{T}},\ldots,\theta_J^{\mathrm{T}},\rho^{\mathrm{T}}\right)^{\mathrm{T}}$ and θ_j's are vectors of parameters of the jth SF regression. Therefore, the ML estimator of θ is defined as

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \ln L(\theta),$$

where Θ denotes the parameter space of θ. Under the regularity conditions for the asymptotic maximum likelihood theory, the ML estimator can be shown to be consistent, asymptotic efficient and asymptotic normal,[4] that is,

$$\sqrt{N}\left(\hat{\theta} - \theta_0\right) \to N\left(0, I^{-1}(\theta_0)\right),$$

where $I(\theta_0)$ is the usual Fisher's information matrix and $\theta_0 = \left(\theta_{1,0},\ldots,\theta_{J,0},\rho\right)$ denotes the vector of true parameters.

The objective function in (8a) and (8b) give us further insight on the difference in the estimators by separate

[3] See Proposition 4.1 of Cherubini et al. (2004).

[4] See Serfling (1980).

estimation and joint estimation. The first term in (8a) captures the correlations between the equations; while the second term is the log-likelihood function obtained by separate regression. Although the separate regression may still give consistent estimates and also valid standard errors under correctly specified marginal densities; however, the standard errors are inefficient. Given the marginal specification, maximizing (8a) gives consistent and more efficient ML estimator if the correct copula density is included.

Empirically, we never know if the copula density is true; therefore, the log-likelihood function may be referred as the log quasi-likelihood function. The derived quasi-maximum likelihood (QML) estimator, which is still consistent (in the sense of the best approximation) but has an invalid standard error. Compared with the separate regression, we still have efficiency gain. However, in this case it is necessary to compute the "sandwich" estimator in order to obtain the correct standard errors.

Some existing tools are helpful in choosing a correctly specified copula. For instance, the goodness of fit tests reviewed in Genest et al. (2009) or the moment test suggested by Amsler et al. (2011). Alternatively, if there are many candidate copula functions, the commonly used model selection criteria, such the Akaike information criterion (AIC) or Bayesian information criterion (BIC) can be also used in deciding the proper copula. See Trivedi and Zimmer (2005) for more discussion on the model selection criteria. For a further consideration, it seems that we are able to implement the Hausman test by comparing the difference in the estimators of separate and joint estimation.[5] If the copula function is misspecified, the two estimators will have different probability limits.

The model specification test devised by Hausman (1987) can be applied to test for the specification of copulas. The essential idea of Hausman (1987) is that the covariance of an efficient estimator with its difference from an inefficient estimator is zero, i.e., $Cov\left(\hat{\theta}_S-\hat{\theta}_J,\hat{\theta}_J\right)=0$, which suggests that

$$Var\left(\hat{\theta}_S-\hat{\theta}_J\right)=Var\left(\hat{\theta}_S\right)-Var\left(\hat{\theta}_J\right),$$

where the subscript S and J denote separate and joint estimation, respectively. Therefore, the Hausman's Chi-squared test statistics is

$$W=\left(\hat{\theta}_S-\hat{\theta}_J\right)^{\mathrm{T}}Var\left(\hat{\theta}_S-\hat{\theta}_J\right)^{-1}\left(\hat{\theta}_S-\hat{\theta}_J\right)\sim\chi^2(k),$$

where k is the number of common parameters in the separate and joint estimation.

Finally, some measures of the dependence structure between the two SF regressions can be obtained by the transformations from the Gaussian copula parameter matrix $\Omega=[\Omega_{js}]$. Taking the transformation according to the distribution function,[6] one can show that the linear correlation between $F_j(\varepsilon_{ji})$ and $F_s(\varepsilon_{si})$ is

$$\rho_{F_j,F_s}=\frac{6}{\pi}\arcsin\left(\frac{\Omega_{js}}{2}\right),$$

which measures the correlation of the two SF regressions (j and s) in terms of the CDF of ε_{ji} and ε_{si} and is also called the Spearman's rank correlation coefficient of ε_{ji} and ε_{si}.

Alternatively, another dependence measure is concordance. For example, two observations $(\varepsilon_{1i_1},\varepsilon_{2i_1})$ and $(\varepsilon_{1i_2},\varepsilon_{2i_2})$ of a pair $(\varepsilon_{1i},\varepsilon_{2i})$ of composite errors are concordant if both values of one pair are greater than the corresponding values of the other pair, that is if $\varepsilon_{1i_1}>\varepsilon_{1i_2}$ and $\varepsilon_{2i_1}>\varepsilon_{2i_2}$, or $\varepsilon_{1i_1}<\varepsilon_{1i_2}$ and $\varepsilon_{2i_1}<\varepsilon_{2i_2}$. The discordance is defined in the opposite way; in other words, ε_{1i} and ε_{2i} are said to be discordant if $\varepsilon_{1i_1}>\varepsilon_{1i_2}$ and $\varepsilon_{2i_1}<\varepsilon_{2i_2}$, or $\varepsilon_{1i_1}<\varepsilon_{1i_2}$ and $\varepsilon_{2i_1}>\varepsilon_{2i_2}$. The measure of the concordance is also called the Kendall's coefficient, which is defined as

$$\begin{aligned}\tau_{12}&=\Pr((\varepsilon_{1i_1}-\varepsilon_{1i_2})(\varepsilon_{2i_1}-\varepsilon_{2i_2})>0)\\&\quad-\Pr((\varepsilon_{1i_1}-\varepsilon_{1i_2})(\varepsilon_{2i_1}-\varepsilon_{2i_2})<0).\end{aligned}$$

Intuitively, Kendall τ measures the difference between the probability of concordance and that of discordance for two composite errors. It can be shown that Kendall τ can also be transformed from the Gaussian copula parameter Ω. The concordance between ε_{ji} and ε_{si} of the two SF regressions (j and s) is

$$\tau_{js}=\frac{2}{\pi}\arcsin(\Omega_{js}).$$

Although we may obtain certain dependence measures of the SF regressions through the direct transformation of the copula parameter Ω, the interpretation of the measure is limited to the composite errors. How much correlation between ε_{ji} and ε_{si} is attributed to (v_{ji}, v_{si}) or (u_{ji}, u_{si}) cannot be further identified.

4 Computation issue

The maximum likelihood estimation in (8) requires the evaluation of the PDF $f_j(\varepsilon_{ji})$ and the inverse (quantile) of the distribution function $F_j(\varepsilon_{ji})$ in the vector $\zeta_i=\left(\Phi^{-1}(F_1(\varepsilon_{1i})),\ldots,\Phi^{-1}(F_J(\varepsilon_{Ji}))\right)^{\mathrm{T}}$. As shown by Aigner et al. (1977), the PDF of the composite error ε_{ji} is

$$f_j(\varepsilon_{ji})=\frac{2}{\sigma_j}\phi\left(\frac{\varepsilon_{ji}}{\sigma_j}\right)\Phi\left(-\frac{\varepsilon_{ji}\lambda_j}{\sigma_j}\right),\tag{9}$$

where $\sigma_j=\sqrt{\sigma_{vj}^2+\sigma_{uj}^2}$, $\lambda_j=\frac{\sigma_{uj}}{\sigma_{vj}}$, and $\phi(.)$ and $\Phi(.)$ are the standard normal PDF and CDF, respectively. The

[5] We thank the associate editor for this suggestion.

[6] See Cherubini et al. (2004) for the details.

computation of the PDF $f_j(\varepsilon_{ji})$ is relatively easy and routine. However, the computation of CDF $F_j(\varepsilon_{ji})$ is not trivial. Since the closed form evaluation of CDF does not exist, it typically requires some numerical integration procedures or the use of simulation to approximate the integration (Greene 2003, 2010) in computing $F_j(\varepsilon_{ji})$. Recently Tsay et al. (2012) have suggested a method utilizing some mathematical approximation functions to obtain the closed form of the CDF of $\varepsilon_{ji} = v_{ji} + u_{ji}$, and shown that the approximation error is less than 10^{-5}. Here, we follow a similar computation procedure and derive the closed form for the case of $\varepsilon_{ji} = v_{ji} - u_{ji}$.

For simplicity of presentation, the subscripts of variables and function are omitted. Given the PDF $f(\varepsilon)$ in (9), the corresponding CDF $F(Q)$ of ε at point Q is

$$\Pr(\varepsilon \le Q) = F(Q) = \frac{2}{\sigma} I(Q), \tag{10}$$

where

$$I(Q) = \int_{-\infty}^{Q} \left(\int_{-\infty}^{-\frac{\varepsilon\lambda}{\sigma}} \phi(\xi) d\xi \right) \phi\left(\frac{\varepsilon}{\sigma}\right) d\varepsilon.$$

It follows that the evaluation of the CDF of ε requires the computation of $I(Q)$.Define a sign function

$$\operatorname{sign}(\xi) = \begin{cases} 1, & \text{if } \xi > 0; \\ 0, & \text{if } \xi = 0; \\ -1, & \text{if } \xi < 0. \end{cases}$$

and an error function

$$\operatorname{erf}(\xi) = 2 \int_0^{\sqrt{2\pi}} \phi(t) dt = 2\Phi(\sqrt{2}\xi) - 1.$$

Let $a = -\frac{\lambda}{\sigma}$, $b = \frac{1}{\sigma}$. Following the approach of Tsay et al. (2009), we may show that the function $I(Q)$ can be approximated by

$$\begin{aligned} I(Q) \approx \frac{1}{2b} &\left[1 + \frac{\operatorname{erf}\left(\frac{bQ}{\sqrt{2}}\right)}{2b} \cdot \left(\frac{1 - \operatorname{sign}(Q)}{2} \right) \right] \\ &- \frac{1}{4\sqrt{b^2 - a^2 c_2}} \exp\left(\frac{a^2 c_1^2}{4(b^2 - a^2 c_2)} \right) \\ &\times \left[1 + \operatorname{erf}\left(\frac{-ac_1 - \sqrt{2}Q(b^2 - a^2 c_2)\operatorname{sign}(Q)}{2\sqrt{b^2 - a^2 c_2}} \right) \right], \end{aligned}$$

where c_1 and c_2 are known constants.[7] Further details about the derivation of (11) can be found in Tsay et al. (2012).

[7] Tsay et al. (2012) have shown that $c_1 = -1.0950081470333$ and $c_2 = -0.75651138383854$.

5 A Monte Carlo simulation

In this section we conduct a Monte Carlo simulation based on a two SF regressions model. The data-generating process (DGP) is set as the following

$$\begin{aligned} &\text{Regression 1 } (F_1): \quad y_1 = \beta_{10} + \beta_{11} x_1 + \varepsilon_1, \\ &\text{Regression 2 } (F_2): \quad y_2 = \beta_{20} + \beta_{21} x_2 + \varepsilon_2, \end{aligned} \tag{12}$$

where $\varepsilon_j = v_j - u_j, j = 1, 2$. The parameters are set as $\beta_{10} = 1$, $\beta_{11} = 1$, $\beta_{20} = 0.5$, and $\beta_{21} = 1.5$. The distributions of the symmetric stochastic errors are assumed to be normal with unequal variances, i.e., $v_j \sim N(0, \sigma_{vj}^2)$, the one-sided errors are assumed to follow the half normal distributions, i.e., $u_j \sim N^+(0, \sigma_{uj}^2)$, and the corresponding transformed parameters specified in (12) are $\sigma_j = \sqrt{\sigma_{vj}^2 + \sigma_{uj}^2}$ and $\lambda_j = \frac{\sigma_{uj}}{\sigma_{vj}}$. In our following simulations, three different settings of the parameters about σ_{uj}^2 and σ_{vj}^2 are specified and summarized below:

Case 1 $\sigma_{v1}^2 = 1.6$, $\sigma_{u1}^2 = 0.4, \sigma_{v2}^2 = 5.55$ and $\sigma_{u2}^2 = 0.5$. The transformed parameters are $\sigma_1 = 1.4142$, $\sigma_2 = 2.4597$, and $\lambda_1 = 0.5$, $\lambda_2 = 0.3$.

Case 2 $\sigma_{v1}^2 = 0.15$, $\sigma_{u1}^2 = 0.4, \sigma_{v2}^2 = 0.2$ and $\sigma_{u2}^2 = 0.5$. The transformed parameters are $\sigma_1 = 0.7416$, $\sigma_2 = 0.8367$, and $\lambda_1 = 1.63$, $\lambda_2 = 1.58$.

Case 3 $\sigma_{v1}^2 = 0.025$, $\sigma_{u1}^2 = 0.4$, $\sigma_{v2}^2 = 0.056$ and $\sigma_{u2}^2 = 0.5$. The transformed parameters are $\sigma_1 = 0.6519$, $\sigma_2 = 0.7457$, and $\lambda_1 = 4$, $\lambda_2 = 2.99$.

Among these cases, we have set, λ_j, the ratios of σ_{uj}andσ_{vj}, increasing from Case 1 to Case 3 in order to demonstrate the effect of λ on the estimates of separate and joint estimation. With known σ_1, σ_2, λ_1, and λ_2, we are able to find $F_1(\varepsilon_1)$ and $F_2(\varepsilon_2)$ for any given ε_1 and ε_2.

We assume that the dependence between the composite errors, ε_1 and ε_2, and hence the dependence between the dependent variables, y_1 and y_2, is captured by the Gaussian copula with the copula parameter ρ. Thus the Gaussian copula representation of the joint CDF of ε_1 and ε_2 is

$$\begin{aligned} F(\varepsilon_1, \varepsilon_2) &= C(F_1(\varepsilon_1), F_2(\varepsilon_2); \rho) \\ &= \Phi_2(\zeta_1, \zeta_2; \rho) \end{aligned}$$

where $\zeta_j = \Phi^{-1}(F_j(\varepsilon_j))$. As shown in (4), ζ_1 and ζ_2 follow a standard bivariate normal distribution with the correlation coefficient ρ. With the copula representation in place, the steps of generating the composite errors ε_1 and ε_2 for the simulation are stated as follows:

Step 1 Draw two independent random variables, z_1 and z_2, from $N(0, 1)$.

Step 2 Compute $\zeta_1 = z_1$ and $\zeta_2 = z_1\rho + z_2\sqrt{1-\rho^2}$ such that ζ_1 and ζ_2 follow a standard bivariate normal distribution with correlation coefficient ρ.

Step 3 Compute $\gamma_1 = \Phi(\zeta_1)$ and $\gamma_2 = \Phi(\zeta_2)$ from normal distribution CDF. γ_1 and γ_2 are the probability integral transformations of ζ_1 and ζ_2.

Step 4 Finally, the composite errors, ε_1 and ε_2, are generated by the inverse CDF, $\varepsilon_1 = F_1^{-1}(\gamma_1)$ and $\varepsilon_2 = F_2^{-1}(\gamma_2)$.[8]

Step 5 Given the parameters β_{10}, β_{11}, β_{20}, and β_{21} and also the exogenous variables x_1 and x_2, the dependent variables y_1 and y_2 are generated from the SF regressions in (12).

It is worth emphasizing that the composite errors ε_1 and ε_2 generated in Step 4 are correlated, and the dependence may implicitly come from either the correlation between v_1 and v_2, or between u_1 and u_2, or both.

In our simulation, the number of replications is set to be 1,000 and the small sample performance is investigated with the sample sizes $N = 200$, 500. Various degrees of dependence structure are tried experimentally with the dependence parameter $\rho = 0.1$, 0.3, 0.5, 0.7, 0.9 considered. Both of the exogenous variables x_1 and x_2 are independently drawn from uniform distributions over the range [1, 2] and have been kept fixed over the replications, only ε_1 and ε_2 are redrawn in each replication.

The performance comparisons of the separate and joint estimators are based on the relative mean squared errors (RMSE). Two sets of estimators are evaluated: the coefficients of the frontier part, (β_{10}, β_{11}, β_{20}, β_{21}), and the variances of error components, (σ_{v1}^2, σ_{u1}^2, σ_{v2}^2, σ_{u2}^2). The relative performance of the estimators are investigated by the RMSEs, which are defined as

$$\mathrm{RMSE}(k) = \frac{\mathrm{MSE}_S(k)}{\mathrm{MSE}_J(k)},$$

where k stands for the regression coefficients or the variances of error components. The subscript '*S*' denotes separate estimation and '*J*' denote joint estimation, respectively. Therefore, $\mathrm{RMSE}(k) > 1$ is equivalent to that $\mathrm{MSE}_S(k) > \mathrm{MSE}_J(k)$. Similarly, we also use the relative efficiency (RE), defined as the relative variance of the estimators,

$$\mathrm{RE}(k) = \frac{\mathrm{Var}_S(k)}{\mathrm{Var}_J(k)},$$

to investigate the efficiency gain. The simulated results of RMSE and RE under Cases 1–3 are summarized in Tables 1 and 2, respectively.

A few findings from Table 1 are worthy of mention. First, Table 1 shows that almost all RMSEs are greater than 1, except for few cases having RMSEs equal to 0.99 when $\rho = 0.1$ and $N = 200$. The joint estimation shows significantly better in the mean squared errors performance than the separate estimation when there is some dependence between the two regressions. The superiority of the joint estimation over the separate estimation shows even more clearly in the estimation of variances of the SF error components, σ_{vj}^2 and σ_{uj}^2 because of the relatively high values of RMSEs. This finding suggests that incorporating the correlation between the SF regressions may significantly increase the efficiency in estimating the variances, and thus the efficiency in predicting the technical efficiency.

Second, for a fixed n and λ, the RMSEs tend to increase as the dependence of the two regressions gets stronger, which suggests that the loss of estimation efficiency will worsen when the dependence between the two sectors increases in strength.

Third, for each fixed ρ and λ, the RMSEs for the β coefficients as well as σ_{vj}^2 and σ_{uj}^2 increase when the sample size increases from 200 to 500, which is consistent with the implications of the large sample theory. Therefore, we may conclude that the overall performance of the estimators under joint estimation is in general better than that under separate estimation. In other words, if the dependence structure is ignored in the estimation process, the loss of estimation efficiency will, as expected, worsen as the dependence between the two sectors becomes stronger or as the sample size increases. Cases 2 and 3 also suggest similar findings as λ, the ratio of σ_u and σ_v, becomes larger in the two regressions.

Finally, by comparing Cases 1–3 we also find that the RMSEs of σ_u and σ_v for all ρ's significantly increase, which indicates that as λ increases the joint estimation also become significantly better in the mean squared errors performance than the separate estimation even for fixed N and ρ.

Similar findings are also found in the REs. Most REs are greater than 1, except for σ_{u1}^2, σ_{u2}^2 and few cases when $n = 200$ and $\rho = 0.1$. The joint estimation indeed shows higher efficiency gain than the separate estimation, especially when the dependence of the two regressions gets stronger or the sample size increases.

In addition to RMSEs and REs, we also investigate the bias of the estimators obtained by separate and joint estimations. The estimated biases of the parameters, β_{10}, β_{11}, σ_{v1}^2, and σ_{u1}^2 in Regression 1 and $\beta_{20}, \beta_{21}, \sigma_{v2}^2$, and σ_{u2}^2 in

[8] Since the close form of the inverse CDF is not available, we first compute, via (10), the numerical CDF at 10,000 points, and then apply the inverse of this numerical CDF at given γ to obtain ε.

Table 1 RMSEs of Cases 1–3 when the sample sizes are 200 and 500

ρ	0.1		0.3		0.5		0.7		0.9	
n	200	500	200	500	200	500	200	500	200	500
Regression 1										
Case 1: ($\lambda_1 = 0.5$)										
β_{10}	1.119	1.190	1.148	1.268	1.309	1.379	1.554	1.544	2.440	2.152
β_{11}	0.995	1.006	1.059	1.096	1.258	1.326	1.809	1.911	4.822	4.985
σ^2_{v1}	11.832	24.215	12.116	24.864	12.623	25.397	13.683	26.804	15.992	32.533
σ^2_{u1}	1.509	3.396	1.537	3.517	1.610	3.629	1.765	3.863	2.238	5.016
Case 2: ($\lambda_1 = 1.63$)										
β_{10}	0.996	1.005	1.027	1.081	1.159	1.267	1.515	1.721	3.034	3.356
β_{11}	0.991	1.001	1.053	1.089	1.247	1.313	1.796	1.901	4.776	4.942
σ^2_{v1}	31.978	71.305	32.679	73.877	34.496	79.995	38.548	93.639	49.149	126.459
σ^2_{u1}	2.276	6.276	2.309	6.474	2.418	6.955	2.685	8.040	3.388	10.656
Case 3: ($\lambda_1 = 4$)										
β_{10}	0.993	1.002	1.044	1.075	1.199	1.256	1.629	1.703	3.617	3.558
β_{11}	0.992	1.003	1.051	1.089	1.230	1.299	1.735	1.844	4.495	4.638
σ^2_{v1}	1,254.80	3,143.32	1,292.51	3,237.64	1,385.15	3,486.19	1,599.57	4,053.26	2,176.97	5,609.69
σ^2_{u1}	29.613	79.433	29.883	80.750	30.654	83.820	32.457	90.095	36.640	103.400
Regression 2										
Case 1: ($\lambda_2 = 0.3$)										
β_{20}	1.160	1.243	1.281	1.359	1.412	1.466	1.730	1.579	2.330	2.054
β_{21}	1.002	1.014	1.072	1.100	1.291	1.327	1.919	1.883	5.317	4.800
σ^2_{v2}	1.431	1.547	1.434	1.552	1.438	1.554	1.439	1.554	1.444	1.557
σ^2_{u2}	1.351	3.037	1.374	3.124	1.439	3.220	1.589	3.444	2.051	4.538
Case 2: ($\lambda_2 = 1.58$)										
β_{20}	1.015	1.016	1.084	1.097	1.247	1.272	1.700	1.657	3.244	3.119
β_{21}	0.993	1.011	1.064	1.103	1.282	1.324	1.902	1.884	5.305	4.718
σ^2_{v2}	13.127	14.494	13.278	14.662	13.621	14.990	14.258	15.567	15.466	16.514
σ^2_{u2}	3.039	5.859	3.076	5.947	3.194	6.168	3.475	6.621	4.173	7.482
Case 3: ($\lambda_2 = 2.99$)										
β_{20}	0.996	1.015	1.057	1.101	1.231	1.289	1.720	1.732	3.980	3.552
β_{21}	0.994	1.012	1.056	1.101	1.251	1.317	1.840	1.863	5.195	4.579
σ^2_{v2}	111.879	119.977	112.379	120.300	113.216	120.732	114.347	121.284	115.548	122.167
σ^2_{u2}	15.631	19.132	15.695	19.171	15.907	19.326	16.400	19.648	17.446	20.206

Regression 2, under Cases 1–3 with sample size $N = 200$, 500 are summarized in Tables 3 and 4.

Tables 3 and 4 do not provide significant evidence showing that estimators of the frontier coefficients under joint estimation are definitely better or worse than those under separate estimation. However, the estimated biases of the variance estimators, σ^2_{vj} and σ^2_{uj}, under the joint estimation are much smaller than those under separate estimation for all ρ's. It seems that estimation without considering the correlation between the SF regressions cause a lot of loss in efficiency in the variance estimators and thus in the prediction of technical efficiencies, but not necessarily for the prediction of the SF frontiers. Furthermore, by definition λ represents the relative contribution of u and v to ε. A larger λ implies that the statistical noise plays a relative small role in the composite error. By comparing Cases 1–3 in Tables 3 and 4, both tables suggest that the degree of estimation bias for σ^2_{vj} and σ^2_{uj} is alleviated as λ increases in both separate and joint estimation. In summary, it is of fundamental importance to take into account the interdependence of the SF regressions from the point of view in reducing estimation bias and improving the estimation efficiency.

Table 2 REs of Cases 1–3 when the sample sizes are 200 and 500

ρ	0.1		0.3		0.5		0.7		0.9	
n	200	500	200	500	200	500	200	500	200	500
Regression 1										
Case 1: ($\lambda_1 = 0.5$)										
β_{10}	1.118	1.187	1.147	1.265	1.308	1.378	1.554	1.542	2.458	2.145
β_{11}	0.995	1.006	1.059	1.094	1.258	1.324	1.810	1.906	4.824	4.969
σ^2_{v1}	6.760	6.965	6.891	7.060	7.125	7.183	7.794	7.603	9.194	9.285
σ^2_{u1}	0.154	0.153	0.156	0.156	0.163	0.161	0.181	0.172	0.237	0.227
Case 2: ($\lambda_1 = 1.63$)										
β_{10}	1.118	1.005	1.023	1.081	1.153	1.268	1.506	1.723	3.008	3.363
β_{11}	0.995	1.001	1.053	1.088	1.247	1.312	1.796	1.897	4.779	4.928
σ^2_{v1}	6.760	10.157	10.858	10.526	11.459	11.400	12.794	13.346	16.277	18.028
σ^2_{u1}	0.154	0.099	0.096	0.102	0.101	0.109	0.112	0.127	0.140	0.168
Case 3: ($\lambda_1 = 4$)										
β_{10}	0.993	1.002	1.043	1.075	1.199	1.256	1.628	1.702	3.618	3.555
β_{11}	0.992	1.003	1.051	1.088	1.230	1.298	1.735	1.841	4.497	4.630
σ^2_{v1}	40.924	39.149	42.148	40.328	45.171	43.426	52.211	50.487	71.413	69.863
σ^2_{u1}	0.025	0.026	0.025	0.026	0.025	0.027	0.027	0.029	0.031	0.033
Regression 2										
Case 1: ($\lambda_2 = 0.3$)										
β_{20}	1.160	1.243	1.273	1.345	1.419	1.459	1.790	1.573	2.623	2.107
β_{21}	1.002	1.015	1.072	1.103	1.292	1.332	1.925	1.890	5.339	4.814
σ^2_{v2}	6.531	7.086	6.752	7.064	6.603	7.398	7.023	7.798	8.547	9.713
σ^2_{u2}	0.161	0.163	0.169	0.172	0.175	0.187	0.200	0.206	0.278	0.278
Case 2: ($\lambda_2 = 1.58$)										
β_{20}	1.160	1.017	1.084	1.100	1.246	1.277	1.696	1.667	3.223	3.142
β_{21}	1.002	1.012	1.065	1.104	1.284	1.327	1.909	1.888	5.335	4.726
σ^2_{v2}	6.531	10.002	9.863	10.289	10.134	10.951	11.203	12.363	14.382	16.834
σ^2_{u2}	0.161	0.100	0.106	0.104	0.116	0.114	0.134	0.131	0.161	0.174
Case 3: ($\lambda_2 = 2.99$)										
β_{20}	0.996	1.016	1.058	1.102	1.233	1.292	1.726	1.735	4.009	3.554
β_{21}	0.994	1.012	1.056	1.102	1.254	1.319	1.847	1.866	5.226	4.581
σ^2_{v2}	22.774	23.550	22.523	23.803	23.131	24.695	25.683	27.911	37.647	39.691
σ^2_{u2}	0.044	0.042	0.047	0.044	0.051	0.046	0.057	0.050	0.069	0.060

6 An empirical example

An empirical application of a model with two related stochastic frontier regressions is applied to a study of the hotel industry in Taiwan, focusing on production efficiency in the accommodation and restaurant divisions. The data come from the annual report of the Taiwan Tourism Bureau at the Ministry of Transportations and Communications. Pooled observations of 50 international grand hotels from 2001 to 2005 has been obtained, providing a total of 250 sample points for the empirical study. The two SF regressions, with each representing the operation of the sub-DMU in the accommodation and in the restaurant divisions of an international hotel, are specified. Although the data under consideration is balanced panel, at the current stage we do not take into account the firm-specific effect, such as the fixed effect, in our following analysis due to the complexity of the panel likelihood function in this current SF model under consideration. We leave it to our future study

The output for the accommodation division is measured in total revenue (y_1), while the inputs include the total number of workers (x_{11}), the total number of rooms (x_{12}), and other expenses (x_{13}), which includes utilities, materials, maintenance fees, and so on. The output and inputs are allocable within the accommodation division. The output for the restaurant division is also measured by the total revenue (y_2) while the corresponding inputs are the total

Table 3 Bias of Regression 1 of Cases 1–3 when the sample sizes are 200 and 500

ρ	0.1		0.3		0.5		0.7		0.9	
n	200	500	200	500	200	500	200	500	200	500
Case 1: ($\lambda_1 = 0.5$)										
$\beta_{10}(S)$	−0.0189	−0.0290	−0.0189	−0.0290	−0.0189	−0.0290	−0.0189	−0.0290	−0.0189	−0.0290
$\beta_{10}(J)$	0.0005	−0.0059	0.0000	−0.0160	0.0003	−0.0194	0.0166	−0.0178	0.0384	−0.0062
$\beta_{11}(S)$	0.0002	−0.0108	0.0002	−0.0108	0.0002	−0.0108	0.0002	−0.0108	0.0002	−0.0108
$\beta_{11}(J)$	−0.0006	−0.0100	−0.0020	−0.0082	−0.0033	−0.0061	−0.0041	−0.0034	−0.0034	−0.0004
$\sigma^2_{v1}(S)$	−0.8637	−1.0150	−0.8637	−1.0150	−0.8637	−1.0150	−0.8637	−1.0150	−0.8637	−1.0150
$\sigma^2_{v1}(J)$	−0.1250	−0.0654	−0.1215	−0.0590	−0.1156	−0.0566	−0.1150	−0.0564	−0.1098	−0.0535
$\sigma^2_{u1}(S)$	1.0673	1.1306	1.0673	1.1306	1.0673	1.1306	1.0673	1.1306	1.0673	1.1306
$\sigma^2_{u1}(J)$	0.3157	0.1742	0.3081	0.1571	0.2952	0.1523	0.2988	0.1540	0.2905	0.1482
Case 2: ($\lambda_1 = 1.63$)										
$\beta_{10}(S)$	−0.0286	−0.0014	−0.0286	−0.0014	−0.0286	−0.0014	−0.0286	−0.0014	−0.0286	−0.0014
$\beta_{10}(J)$	−0.0272	−0.0015	−0.0251	−0.0020	−0.0223	−0.0026	−0.0182	−0.0031	−0.0110	−0.0034
$\beta_{11}(S)$	0.0005	−0.0042	0.0005	−0.0042	0.0005	−0.0042	0.0005	−0.0042	0.0005	−0.0042
$\beta_{11}(J)$	0.0002	−0.0040	−0.0007	−0.0035	−0.0014	−0.0026	−0.0017	−0.0015	−0.0014	−0.0002
$\sigma^2_{v1}(S)$	0.2370	0.2456	0.2370	0.2456	0.2370	0.2456	0.2370	0.2456	0.2370	0.2456
$\sigma^2_{v1}(J)$	0.0035	0.0016	0.0035	0.0016	0.0033	0.0016	0.0028	0.0015	0.0015	0.0014
$\sigma^2_{u1}(S)$	−0.2464	−0.2484	−0.2464	−0.2484	−0.2464	−0.2484	−0.2464	−0.2484	−0.2464	−0.2484
$\sigma^2_{u1}(J)$	−0.0126	−0.0045	−0.0122	−0.0045	−0.0110	−0.0041	−0.0085	−0.0034	−0.0035	−0.0024
Case 3: ($\lambda_1 = 4$)										
$\beta_{10}(S)$	−0.0027	0.0023	−0.0027	0.0023	−0.0027	0.0023	−0.0027	0.0023	−0.0027	0.0023
$\beta_{10}(J)$	−0.0023	0.0022	−0.0016	0.0018	−0.0007	0.0013	0.0004	0.0007	0.0018	0.0001
$\beta_{11}(S)$	0.0006	−0.0021	0.0006	−0.0021	0.0006	−0.0021	0.0006	−0.0021	0.0006	−0.0021
$\beta_{11}(J)$	0.0004	−0.0020	0.0000	−0.0018	−0.0004	−0.0013	−0.0007	−0.0007	−0.0007	1.0000
$\sigma^2_{v1}(S)$	0.3764	0.3749	0.3764	0.3749	0.3764	0.3749	0.3764	0.3749	0.3764	0.3749
$\sigma^2_{v1}(J)$	−0.0005	0.0001	−0.0005	0.0001	−0.0005	0.0001	−0.0005	0.0001	−0.0007	0.0001
$\sigma^2_{u1}(S)$	−0.3755	−0.3749	−0.3755	−0.3749	−0.3755	−0.3749	−0.3755	−0.3749	−0.3755	−0.3749
$\sigma^2_{u1}(J)$	0.0015	−0.0001	0.0016	−0.0001	0.0019	2*E-5	0.0028	0.0004	0.0044	0.0010

S and *J* in the parentheses denote separate and joint estimation, respectively

number of workers (x_{21}), the floor area of the restaurant (x_{22}), and other expenses (x_{23}), including utilities, materials, and so on. Again, these output and inputs are attributable and accountable within the restaurant division. All revenues and other expenses are measured in New Taiwan dollars (NT$). Since these two divisions of a hotel share certain common characteristics, such as the same DMU, brand, location… etc.; the outputs y_1 and y_2 should be correlated and so are the composite errors. The stochastic frontier regressions in accommodation and restaurant of a hotel operation are specified as:

$$\text{Accommodation: } y_{1i} = \beta_{10} + \beta_{11}x_{11i} + \beta_{12}x_{12i} + \beta_{13}x_{13i} + \varepsilon_{1i},$$
$$\text{Restaurant: } y_{2i} = \beta_{20} + \beta_{21}x_{21i} + \beta_{22}x_{22i} + \beta_{23}x_{23i} + \varepsilon_{2i},$$

where $\varepsilon_{ji} = v_{ji} - u_{ji}$ with the assumption of normal-half normal distribution specification, $v_{ji} \sim N(0, \sigma^2_{vj})$ and $u_{ji} \sim N^+(0, \sigma^2_{uj})$.

The ML estimates of the two SF regression models, using both separate and joint estimations, are summarized in Table 5. The separate estimation results are obtained by estimating the frontier regression in isolation without consideration of the potential dependence between the two divisions. The joint ML estimation is conducted with the dependence of the two regressions specified as the Gaussian copula with the PDF $f(\varepsilon_{1i}, \varepsilon_{2i})$ shown in (4).

The estimated model, either by the separate or the joint ML estimation shown in Table 5, gives no surprising results. All estimates have expected signs and, with the exception of the coefficient of the floor area of restaurant (x_{22}), are statistically significant at the 5 % level or better. It is observed in Table 5 that the point estimates yielded by the separate and joint methods differ. This outcome is to be expected since different likelihood functions in (8) are maximized in two methods, unless the dependence parameter $\rho = 0$. The estimated dependence in the

Table 4 Bias of Regression 2 of Cases 1–3 when the sample sizes are 200 and 500

ρ	0.1		0.3		0.5		0.7		0.9	
n	200	500	200	500	200	500	200	500	200	500
Case 1: ($\lambda_2 = 0.3$)										
$\beta_{20}(S)$	0.3351	0.1296	0.3467	0.1403	0.3438	0.1380	0.3339	0.1564	0.2808	0.1921
$\beta_{20}(J)$	0.3111	0.1174	0.2953	0.0866	0.2969	0.1016	0.3040	0.1149	0.3250	0.1669
$\beta_{21}(S)$	−0.0417	0.0161	−0.0412	0.0117	−0.0378	0.0062	−0.0304	−0.0009	−0.0151	−0.0105
$\beta_{21}(J)$	−0.0391	0.0209	−0.0383	0.0216	−0.0368	0.0202	−0.0303	0.0167	−0.0168	0.0103
$\sigma^2_{v2}(S)$	−3.3230	−3.9268	−3.2984	−3.9277	−3.2906	−3.9516	−3.2868	−3.9495	−3.3468	−3.9040
$\sigma^2_{v2}(J)$	−0.5811	−0.3865	−0.5639	−0.3468	−0.5651	−0.3480	−0.5568	−0.3357	−0.5417	−0.3306
$\sigma^2_{u2}(S)$	4.3946	4.6013	4.3785	4.6008	4.3683	4.6116	4.3602	4.6166	4.3808	4.6103
$\sigma^2_{u2}(J)$	1.5238	0.9517	1.4615	0.8422	1.4579	0.8566	1.4353	0.8455	1.4113	0.8663
Case 2: ($\lambda_2 = 1.58$)										
$\beta_{20}(S)$	−0.0136	−0.0159	−0.0122	−0.0149	−0.0120	−0.0132	−0.0154	−0.0100	−0.0233	−0.0050
$\beta_{20}(J)$	−0.0119	−0.0167	−0.0104	−0.0164	−0.0081	−0.0152	−0.0062	−0.0127	−0.0052	−0.0082
$\beta_{21}(S)$	−0.0109	0.0035	−0.0109	0.0024	−0.0101	0.0010	−0.0079	−0.0006	−0.0038	−0.0029
$\beta_{21}(J)$	−0.0112	0.0042	−0.0111	0.0041	−0.0104	0.0038	−0.0086	0.0033	−0.0049	0.0021
$\sigma^2_{v2}(S)$	0.2898	0.2912	0.2914	0.2905	0.2925	0.2902	0.2915	0.2912	0.2885	0.2930
$\sigma^2_{v2}(J)$	0.0032	0.0011	0.0025	0.0012	0.0016	0.0013	0.0009	0.0014	0.0002	0.0014
$\sigma^2_{u2}(S)$	−0.2968	−0.2988	−0.2978	−0.2986	−0.2986	−0.2984	−0.2987	−0.2984	−0.2979	−0.2983
$\sigma^2_{u2}(J)$	−0.0100	−0.0086	−0.0087	−0.0088	−0.0067	−0.0084	−0.0044	−0.0069	−0.0012	−0.0042
Case 3: ($\lambda_2 = 2.99$)										
$\beta_{10}(S)$	0.0095	−0.0044	0.0094	−0.0033	0.0083	−0.0018	0.0060	0.0002	0.0014	0.0030
$\beta_{10}(J)$	0.0098	−0.0050	0.0100	−0.0049	0.0095	−0.0045	0.0087	−0.0037	0.0070	−0.0021
$\beta_{11}(S)$	−0.0085	0.0018	−0.0084	0.0010	−0.0074	0.0001	−0.0060	−0.0010	−0.0033	−0.0027
$\beta_{11}(J)$	−0.0087	0.0022	−0.0086	0.0022	−0.0077	0.0022	−0.0063	0.0020	−0.0038	0.0014
$\sigma^2_{v1}(S)$	0.4458	0.4423	0.4459	0.4422	0.4464	0.4424	0.4459	0.4429	0.4451	0.4435
$\sigma^2_{v1}(J)$	−0.0003	−0.0004	−0.0006	−0.0004	−0.0009	−0.0004	−0.0013	−0.0003	−0.0017	−0.0001
$\sigma^2_{u1}(S)$	−0.4444	−0.4444	−0.4446	−0.4444	−0.4449	−0.4444	−0.4450	−0.4443	−0.4449	−0.4441
$\sigma^2_{u1}(J)$	0.0018	−0.0016	0.0022	−0.0016	0.0031	−0.0011	0.0048	−0.0002	0.0074	0.0011

S and *J* in the parentheses denote separate and joint estimation, respectively

Gaussian copula $\hat{\rho} = 0.6318$ is statistically significant at the 1 % level. The corresponding Spearman's rank correlation coefficient is 0.7284, which provides evidence of a significant positive rank correlation between ε_1 and ε_2. The estimated Kendall's τ is about 0.5345.

For the current case, the separate estimation is equivalent to joint estimation with the restriction $\rho = 0$ or using the product copula. Therefore, we adapt the likelihood ratio (LR) statistics for testing the hypothesis H_0: $\rho = 0$. The value of the LR statistics is 103.83 so the Gaussian copula is suggested. Moreover, both model selection criteria AIC and BIC consistently suggest the Gaussian copula.

It is also seen from the results in Table 5 that the application of the joint estimation has resulted in a significant reduction (about 12 %) in the standard errors of the estimated frontier coefficients $\hat{\beta}$ as compared with those of a separate regression-by-regression estimation. Even if the separate ML estimation provides a consistent estimator of the parameter, it does not fully incorporate all the information relating to the correlation between the two divisions; and thus, the estimates are not as efficient as those obtained under the joint estimation method. More evidence is provided through a comparison of the estimated coefficient of x_{22}, which is insignificant in the separate estimation, and yet significant in the joint estimation at less than the 1 percent significance level.

For rigorousness, we further conduct the Hausman test to test for the specification of copula function. Unfortunately, the test statistics is 11.518787, the Gaussian copula is rejected. Together with the previous LR test result, it seems that the two divisions have significant dependence; and the Gaussian copula can give a better approximation of the true joint density than the product copula (or separate estimation) but may not represent the true one. Therefore, the QML standard errors should be considered for further statistical inferences.

Table 5 The separate and joint ML estimation results

Regression	Variables	Coefficients	
		Separate estimation	Joint estimation
Accommodation: y_1	x_{11}	0.4139 (0.0458)***	0.2750 (0.0400)*** [0.0532]***
	x_{12}	0.1002 (0.0501)**	0.1842 (0.0386)*** [0.0483]***
	x_{13}	0.6034 (0.0271)***	0.6311 (0.0230)*** [0.0338]***
	Constant	6.1910 (0.4091)***	5.7665 (0.3593)*** [0.5131]***
	σ_v^2	0.0309 (0.0124)***	0.0440 (0.0154)*** [0.0153]***
	σ_u^2	0.0507 (0.0360)***	0.0189 (0.0413)*** [0.0398]***
Estimated TE	Mean	0.84252	0.89942
	s.d.	0.06702	0.02670
	min	0.62137	0.79716
	max	0.94887	0.94951
Restaurant: y_2	x_{21}	0.3129 (0.0308)***	0.2574 (0.0302)*** [0.0403]***
	x_{22}	0.0313 (0.0216)	0.0696 (0.0191)*** [0.0154]***
	x_{23}	0.7755 (0.0335)***	0.7575 (0.0300)*** [0.0409]***
	Constant	3.1455 (0.4452)***	3.4775 (0.4043)*** [0. 5306]***
	σ_v^2	0.0059 (0.0018)***	0.0083 (0.0022)*** [0.0023]***
	σ_u^2	0.0836 (0.0112)***	0.0800 (0.0114)*** [0.0119]***
Estimated TE	Mean	0.80813	0.81200
	s.d.	0.12038	0.11335
	min	0.43599	0.43755
	max	0.97263	0.96527
Copula parameter	ρ	–	0.6318 (0.0089)*** [0.0138]***
Log-likelihood	Accommodation	22.347	
	Restaurant	74.115	
	Total	96.462	148.379
# of parameters		12	13
Model selection	AIC	−168.924	−270.7582
	BIC	−126.667	−224.9793
Correlation	ρ_{F_1,F_2}	–	0.7284 (0.0647)*** [0.1004]***
	τ_{12}	–	0.5345 (0.0310)*** [0.0482]***

The Chi-squared test statistic W = 11.52 > $\chi^2_{0.05}(12) = 5.23$. LR test statistics = 103.83426 > $\chi^2_{0.05}(1) = 0.00393$

*, **, and *** represent p-values less than 0.1, 0.05, and 0.01, respectively

Numbers in the parentheses and brackets are the ML and QML (using the "sandwich" formula) standard errors, respectively

With the ML estimates of the stochastic frontier regressions, we further compare the estimates of the individual hotel's technical efficiency based on the separated and joint methods. Following Battese and Coelli (1988), we define the technical efficiency of the jth division of ith hotel as the conditional expectation, $\mathrm{TE}_{ji} = \mathrm{E}\left(e^{-u_{ji}}|\varepsilon_{ji}\right)$. Under the normal-half normal distribution assumption, the estimated technical efficiency is

$$\widehat{\mathrm{TE}}_{ji} = \frac{1-\Phi\left(\hat{\sigma}_{*j} - \hat{\mu}_{*ji}/\hat{\sigma}_{*j}\right)}{1-\Phi\left(-\hat{\mu}_{*ji}/\hat{\sigma}_{*j}\right)} e^{-\hat{\mu}_{*ji}+\frac{1}{2}\hat{\sigma}^2_{*j}},$$

where $\hat{\mu}_{*ji} = -\hat{\varepsilon}_{ji}\hat{\sigma}^2_{uj}/\hat{\sigma}^2_j$ and $\hat{\sigma}^2_{*j} = \hat{\sigma}^2_{uj}\hat{\sigma}^2_{vj}/\hat{\sigma}^2_j$. We emphasize that, in the joint estimation, the prediction $\hat{\varepsilon}_{ji} = y_{ji} - x^{\mathrm{T}}_{ji}\hat{\beta}_j$ should represent the predicted composite errors calculated from the joint estimates of β; and in the separated estimation, the predicted composite errors should be those calculated from the regression-by-regression estimates of β. Obviously, these two predicted composite errors calculated from the joint and separated estimations are not the same unless the estimated dependence $\hat{\rho}$ is zero.

The summary statistics of the technical efficiency levels of the two divisions are given in Table 5, which shows that, under the separate estimation method, the sample mean predicted technical efficiencies in both divisions are consistently smaller than those obtained under the joint estimation method. This result seems to suggest that the efficiency levels are underestimated in the separate estimation method as a result of the failure to consider the

Table 6 The correlation coefficients of the predicted technical efficiencies under separate and joint estimations

	$TE_1(S)$	$TE_2(S)$	$TE_1(J)$	$TE_2(J)$
$TE_1(S)$	1.0000	0.5417	0.9751	–
$TE_2(S)$	0.5417	1.0000	–	0.9810
$TE_1(J)$	0.9751	–	1.0000	0.6158
$TE_2(J)$	–	0.9810	0.6158	1.0000

TE_1 and TE_2 denote the predicted technical efficiencies of the accommodation and the restaurant division, respectively. S and J in the parentheses denote separate and joint estimation, respectively

potential dependency between the two divisions within the same hotel management.

Table 6 shows the sample correlation coefficients to technical efficiency between the two divisions from the joint and separated methods. The joint method shows a much higher correlation coefficient between the two divisions than the counterpart under the separate method, 0.6158 versus 0.5417. Although the magnitude of technical efficiency measures varies by method, the ordering of the estimated individual technical efficiency, as indicated by the correlation between two methods, is fairly high, 0.9751 in the accommodation division and 0.9810 in the restaurant division. However, it seems that the separate regression-by-regression method is likely to underestimate the joint-interrelated hotel operations. What it is important to realize is that it makes good sense to use the joint method in estimating the seemingly unrelated stochastic frontier regressions.

7 Conclusions

In this paper, we have presented a copula-based maximum likelihood approach to estimate the multiple stochastic frontier regressions with correlated composite errors. The joint estimation of the multiple SF regressions is more efficient than the separate regression-by-regression estimation, as the joint approach takes into consideration the correlation among composite errors. The rationale behind the joint estimation of SF regressions is analogous to the classical generalization to system of seemingly unrelated regressions.

Our simulation of a model with two related stochastic frontier regressions demonstrates that joint estimation provides relatively more efficient estimators for parameters than those in the separate estimation. Moreover, we also found that ignoring the dependence between regressions will also cause severe bias in estimating the technical efficiencies; however, it may not have similar effects on the frontier prediction or parameters in the frontier part. In those cases where there are correlations between the different stochastic frontier regressions, it is of fundamental importance for the dependence structure to be taken into consideration in the model estimation. Finally, we conduct an empirical study of the hotel management in accommodation and in restaurant division in Taiwan. Our results show that both the frontier estimates and predicted technical efficiencies differ significantly if the dependence structure is ignored. In particular, the predicted individual technical efficiencies may be underestimated for both regressions in separate estimation.

Following a similar procedure, the multiple stochastic frontier regressions of microeconomic temporal cross-section data can be extended to macroeconomic settings where the individual regressions are countries within the same economic region or block. The Gaussian copula-based approach of estimating a system of stochastic frontier regressions proposed in this paper can also be further modified to various measures of dependence structures among the regressions by incorporating different copula functions. We leave such analysis to future study.

Acknowledgments Lai gratefully acknowledges the National Science Council of Taiwan (NSC-98-2410-H-194-035) for the research support. The authors thank an anonymous referee and the associate editor for helpful comments.

References

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

Amsler C, Prokhorov A, Schmidt P (2011) Using copulas to model time dependence in stochastic frontier models. Econom Rev (forthcoming)

Bandyopadhyay D, Das A (2006) On measures of technical inefficiency and production uncertainty in stochastic frontier production model with correlated error components. J Prod Anal 26:165–180

Battese G, Coelli T (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. J Econom 38:387–399

Carta, Steel (2010) Modelling multi-output stochastic frontiers using copulas. Comput Statist Data Anal (forthcoming)

Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. Wiley, New York

Genest C, Remillard B, Beausoin D (2009) Goodness-of-fit tests for copulas: a review and a power study. Insurance: Math Econom 44:199–213

Greene WH (2003) Simulated likelihood estimation of the normal-gamma stochastic frontier function. J Prod Anal 19:179–190

Greene WH (2010) A stochastic frontier model with correction for sample selection. J Prod Anal 34:15–24

Hausman JA (1987) Specification tests in econometrics. Econometrica 46(6):1251–1271

Pal M (2004) A note on a unified approach to the frontier production function models with correlated non-normal error components: the case of cross-section data. Indian Econ Rev 39:7–18

Pal M, Sengupta A (1999) A model of FPF with correlated error components: an application to Indian agriculture. Sankhyā: Indian J Statis 61:337–350

Pitt MM, Lee L-F (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. J Dev Econ 9: 43–64

Schweizer B, Sklar A (1983) Probability metric spaces. Elsevier Science, New York

Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York

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

Sklar A (1959) Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l'Institut de Statistique de L'Université de Paris 8:229–231

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

Trivedi PK, Zimmer DM (2005) Copula modeling: an introduction for practitioners. Found Trends Econom 1(1):1–111

Tsay WJ, Huang CJ, Fu TT, Ho IL (2012) A simple closed-form approximation for the cumulative distribution function of the composite error of stochastic frontier models. Journal of Productivity Analysis, forthcoming

Wang H-J, Ho C-W (2010) Estimating panel stochastic frontier models with fixed effects by model transformation. J Econom 157(2):289–296

Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Statist Assoc 57:348–368