A simultaneous stochastic frontier model with dependent error components and dependent composite errors: an application to Chinese banking industry

Article

Published Version

Article

A Simultaneous Stochastic Frontier Model with Dependent Error Components and Dependent Composite Errors: An Application to Chinese Banking Industry

Jianxu Liu [1,2], **Mengjiao Wang** [2,*], **Ji Ma** [2,3], **Sanzidur Rahman** [1,4] **and Songsak Sriboonchitta** [2]

[1] Faculty of Economics, Shandong University of Finance and Economics, Jinan 250000, China; liujianxu1984@163.com (J.L.); srahman@plymouth.ac.uk (S.R.)
[2] Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand; majiyn@hotmail.com (J.M.); songsakecon@gmail.com (S.S.)
[3] Institute of Economics, Yunnan Academy of Social Sciences, Kunming 650031, China
[4] Plymouth Business School, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK
[*] Correspondence: mengjiaow1991@163.com

![check for updates]

Abstract: The paper develops a simultaneous equations stochastic frontier model (SFM) with dependent random noise and inefficiency components of individual equations as well as allowing dependence across all equations of the model using copula functions. First, feasibility of our developed model was verified via two simulation studies. Then the model was applied to assess cost efficiency and market power of the banking industry of China using a panel data of 37 banks covering the period 2013–2018. Results confirmed that our simultaneous SFM with dependent random noise and inefficiency components outperformed its predecessor, which is a simultaneous SFM with dependent composite errors but with independent random noise and inefficiency components of individual SFMs as well as the conventional single-equation SFM. Apart from the statistical and computational superiority of our developed model, we also see that Chinese banks in general have a high level of cost efficiency and that competition in the banking industry of China mainly exists in state-owned banks and joint stock banks. Presence of economies of scales as well as diseconomies of scales were found in different banks. Also, the state-owned banks embraced most sophisticated technologies thereby allowing them to operate with the highest level of cost efficiency.

1. Introduction

The stochastic frontier model (SFM) reflects the functional relationship between inputs and outputs produced by enterprises under a given level of technology. For a production function, the SFM is used to calculate efficiency of the production technology by measuring the gap between actual output and the maximum potential output given technology and input levels. For a cost function, the cost efficiency of an enterprise can be calculated by comparing the difference between actual cost and the potential minimum cost given technology, input and output prices. Conventional SFM decomposes the composite error into two components—statistical noise and inefficiency—thereby allowing separation of inefficiency from external random shocks or measurement errors and thus avoiding overestimation of inefficiency [1]. Given this advantage of SFM, it has been applied to a wide variety of research fields to estimate efficiency of firms and/or economic agents, including the banking and finance

sectors [2–4], the agricultural production sector [5,6] as well as the energy industry and environmental performance [7,8]. In recent decades, SFM has been improved in its estimation approaches ranging from semi-parametric and nonparametric estimations, various assumptions regarding marginal distribution and so on. However, these extensions of single or univariate SFM cannot deal with multiple outputs. Also, neglecting dependence among multiple outputs may lead to biased estimates of technical efficiencies [9]. Therefore, when the nature of outputs is multiple and potentially correlated, it is meaningful to apply simultaneous SFMs.

Existing literature does extend the single-equation SFM to solve the issues of multiple outputs or construct multiple SFMs in some ways. For example, Fernández et al. [10] combined parametric aggregator of outputs with the single-equation SFM to deal with multiple outputs. It is clear that this transformation of multiple outputs is a way out to solve multi-output problems, but it is still constrained by the application of a single-equation model and aggregation of multiple outputs could be an issue on its own. Later, Ferreira and Steel [11] applied a multivariate skewed distribution to model the skewness of composed error terms in multi-output equations of stochastic production frontiers. Carta and Steel [12] first proposed a multi-output SFM using copula functions to link inefficiency error terms. Afterwards, Lai and Huang [9] proposed a multiple SFM with correlated composite errors using copula functions and used the maximum likelihood estimation procedure to draw inference from the model. They showed that omission of dependence between composite errors could result in severely biased estimation of technical efficiency. Hereafter, Huang et al. [13] applied the copula-based simultaneous SFM to measure cost efficiency and Lerner index for Russia, Czech Republic, and Poland, etc. In addition, Huang et al. [14] employed the approach to measure competition, innovation and efficiency in Taiwan's banking industry.

Through the above-mentioned literature, we can find that scholars have successfully extended single-equation SFMs to simultaneous SFMs. Moreover, these studies focused on modelling dependence between either the composite errors or the inefficiency terms of equations in simultaneous SFMs, all of which were proved to be more effective than the single-equation SFM. However, as in the case of conventional SFM, simultaneous SFMs were also developed based on the assumption that the statistical noise and inefficiency components are independent. In fact, dependence between the two error components of SFM should not be ignored. The logic behind this argument is that the correlation between statistical noise and inefficiency may arise from factors that are beyond the control of firms, on efficiency.

Some studies have showed that the relaxation of the restrictive independence assumption of the statistical noise and inefficiency components can remarkably improve performance of the conventional SFM. For example, Bandyopadhyay and Das [15] developed a SFM in which the error components were jointly distributed as a truncated bivariate normal, given the condition that the distribution of observational error is negatively skewed. Soon after, Smith [16] proposed a copula-based SFM with dependent error components relaxing the statistical noise-inefficiency independence assumption. Following that, many scholars have applied the copula-based SFM to assess efficiency of decision making units in various fields, such as, applications to analyze technical efficiencies of Moroccan municipalities [17] and intercrop coffee production in northern Thailand [18]. Further, Sriboonchitta et al. [19] proposed a double-copula SFM with sample selection, extending the standard SFM with sample selection by modelling dependent error components using copula functions. Thus, it is important to allow for the dependence not only between composite errors across multiple equations but also between the statistical noise (random error) and inefficiency components in the simultaneous SFM framework. In view of this, we propose a simultaneous SFM with dependent error components of each equation as well as correlated composite errors across equations, which is not currently available in the existing literature. We apply our proposed model to an empirical panel data set of 37 Chinese banks covering the period 2013–2018, to measure market power, cost efficiency, and technology gap ratio of the Chinese banking industry.

Literature abounds on the efficiency analysis of the Chinese banking industry, ranging from examination of cost, production and/or profit efficiency from different perspectives. Among these, some studies investigated interrelationships between efficiency and market power (competition) of Chinese banks but with contrasting evidences. For instance, Lin et al. [20] found that the competition from foreign banks promote efficiency of domestic banks, while Fungáčová et al. [21] concluded that the increase of competition has no significant relationship with the efficiency of Chinese banks. Other studies mainly focused on the influence of financial reforms [22], risk preference [23,24] and bank ownership types [25,26] on efficiency of Chinese banks. Evidence showed that financial reforms and risk taking were both key determinants that affected efficiency of Chinese banks. Nevertheless, the conclusions on the relationship between efficiency and ownership types of banks are rather mixed. For instance, Berger et al. [25] and Fungáčová et al. [27] provided evidence that the state-owned banks suffered from the lowest efficiency compared with other banks, while Chen et al. [28] argued that Chinese state-owned banks were more efficient than other bank groups. Besides, Wang et al. [29] concluded that there is no significant difference in efficiency for banks of different ownership types.

Most of these studies, along with others, mainly applied SFM and/or data envelopment analysis (DEA) and their extensions to analyze efficiency of Chinese banks. Chen et al. [26], Jiang and He [30] and Zhu et al. [24] measured efficiency of Chinese banks using the DEA framework, combined DEA with support vector machines, Malmquist index, and multi-directions efficiency analysis methods, respectively. SFM has also been widely applied and extended to estimate efficiency of the Chinese banking industry, such as, Lin et al. [20], Yin et al. [31] and Fungáčová et al. [27]. Besides, Silva et al. [32] compared SFA and DEA approaches on the efficiency analysis of Chinese banks and concluded that SFA and DEA provided consistent results overall for industry but not for individual banks. Previous researches have provided a systematic and comprehensive analysis on the efficiency of the Chinese banking industry with respect to different research interests and methodologies. However, the application of simultaneous SFM to the Chinese banking industry is quite limited. Huang et al. [33] developed a stochastic network model to assess the efficiency of Chinese banks under production of multistage processes, with the help of copula methods. To the best of our knowledge, there is no research applying simultaneous SFM with dependent error components to analyze efficiency of the Chinese banking industry.

Therefore, the specific contribution of this study to related literature on SFM developments and applications is threefold: First, we develop a simultaneous SFM with dependent error components and dependent composite errors to measure efficiency, which not only allows for dependence between the composite errors of seemingly unrelated stochastic frontier functions, but also captures the dependence between the random noise and inefficiency components within each equation. This is the novelty of our model development which circumvents the limitations in the existing literature on simultaneous SFM. Second, we verify the reasonability and feasibility of relaxing the restrictive assumption of independence between statistical noise and inefficiency components in individual SFM and simultaneous SFM by conducting two simulation studies thereby providing evidence about the consequences of ignoring correlations between the error components. Third, our model is applied to an empirical panel data set of 37 Chinese banks covering the period 2013–2018, in order to measure market power, cost efficiency, scale economy, and technology gap ratio of the Chinese banking industry, which in turn contributes to the relatively limited literature on such analysis for Chinese banks for current years.

The remaining sections of this paper is arranged as follows: In Section 2, we present the basic theories of copula functions, copula-based SFM and then establish the simultaneous SFM with dependent error components. Section 3 describes the detailed process and main findings of the two simulation studies. In Section 4, we apply our proposed model to a panel data of Chinese banks and summarize the empirical results. Section 5 draws conclusions of this study.

2. Methodology

In this section, the basic concepts of copulas are introduced first. Then, we review the theoretical foundations of copula-based SFM. On this basis, we propose a SFM with dependent error components, which not only allows for dependence between the composite errors of two stochastic frontier functions but also captures the dependence between random noise and inefficiency components within each equation.

2.1. Copula Functions

The concept of copula originated from Sklar's theorem. A copula joins univariate distribution functions of random variables to form multivariate (joint) distribution functions to describe the dependence structure among variables [34,35]. Kreinovich et al. [36] mentioned that the copula is the most efficient way of representing multidimensional distributions and, thus, has been successfully applied to many applications in statistics. A bivariate copula is a cumulative distribution function (CDF) of two random variables with uniform margins $[0, 1]$ and support contained in $[0, 1]^2$ [37]. A copula function can be expressed in terms of a joint distribution function H of two random variables X and Y, such that

$$C(u_1, u_2) = H\big(F^{-1}(u_1), G^{-1}(u_2)\big),\tag{1}$$

and

$$H(x, y) = C(F(x), G(y)),\tag{2}$$

where $C(\cdot, \cdot)$ is the copula function, $u_1, u_2 \in [0, 1]$ are the uniform margins, $F(\cdot)$ and $G(\cdot)$ are the continuous marginal distribution functions of X and Y, and $F^{-1}(\cdot)$ and $G^{-1}(\cdot)$ are the corresponding quantile functions [38–40].

The joint probability density function (PDF) of X and Y is factorized as

$$f(x, y) = \frac{\partial^2 H(x, y)}{\partial x \partial y} = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} \cdot \frac{\partial F(x)}{\partial x} \cdot \frac{\partial G(y)}{\partial y} = c(u_1, u_2) \cdot f(x) \cdot g(y),\tag{3}$$

where $f(x)$ and $g(y)$ are the marginal densities of X and Y and $c(\cdot, \cdot)$ is the PDF of the copula distribution.

Many copula families have been developed to model dependence between variables, where different copula families model dependence in different ways. Elliptical copulas and Archimedean copulas are two parametric copula families. Elliptical copulas, such as the Gaussian copula and student t copula, do not have closed-form expressions and are radial symmetric. On the contrary, Archimedean copulas, including Frank, Gumbel, Clayton, and Joe copulas, admit explicit formulas and have simpler forms. Detailed expressions of the commonly used copula families can be found in Sriboonchitta et al. [41] and Wiboonpongse et al. [18].

Different copulas have different ranges of parameters and, so, the degree of dependence modelled by different copulas cannot be compared directly by values of copula parameters. Instead, we may extract Kendall's tau coefficient from the copula functions to compare correlations. Kendall's tau coefficient is measured by the difference between the probability of concordance and the probability of discordance of two pairs of random variables [42]. The Kendall's tau (τ) for the random vector $(X, Y)^T$ is defined as

$$\tau(X, Y) = P[(X - X')(Y - Y') > 0] - P[(X - X')(Y - Y') < 0],\tag{4}$$

where $(X', Y')^T$ is an independent copy of $(X, Y)^T$. Kendall's tau τ is expressible in terms of a copula function:

$$\tau(X, Y) = 4 \iint_{I^2} C(u_1, u_2)\, dC(u_1, u_2) - 1 = 1 - 4 \iint_{I^2} \frac{\partial C^2(u_1, u_2)}{\partial u_1 \partial u_2} du_1 du_2,\tag{5}$$

where u_1 and u_2 are values of the uniform margins.

2.2. Copula-Based Stochastic Frontier Model

A SFM breaks down the composite error into two components: a normally distributed random error term V, which takes into account uncontrollable exogenous factors, and a non-negative error term W, which represents a firm's technical inefficiency [43]. The two error components are assumed to be independent in the conventional SFM. Nevertheless, this assumption of independence can be relaxed by applying copula functions to model the dependence between V and W [16,18], which is the basis of the so-called copula-based SFM. The basic form of a copula-based SFM is given by

$$logY = X\beta + \varepsilon, \tag{6}$$

with

$$\varepsilon = V \pm W, \tag{7}$$

and

$$H(w, v) = C(F_W(w), F_V(v); \theta), \tag{8}$$

where the output Y is positively valued, $X(k \times 1)$ is a vector of regressors, and $\beta(k \times 1)$ is a vector of unknown parameters. The composite error ε contains two components: a symmetric noise V, which is typically assumed to be normally distributed with $\mathcal{N}\left(0, \sigma_V^2\right)$, and a non-negative inefficiency term W, which is usually supposed to be gamma, half-normal or exponentially distributed. For a stochastic production function, the error term has the specification $\varepsilon = V - W$ while, for a stochastic frontier cost function, the error term is specified by $\varepsilon = V + W$ instead [44]. $H(w, v)$ is the joint CDF of W and V modeled by a copula function $C(\cdot, \cdot)$ with the copula parameter θ, $F_V(v)$ and $F_W(w)$ represent the CDFs of $V = v$ and $W = w$, respectively.

The copula-based SFM is reduced to the conventional SFM when W and V are independent, such that

$$C(F_W(w), F_V(v)) = F_W(w) \times F_V(v), \tag{9}$$

while, if W and V are dependent, the copula-based SFM is also referred to as the SFM with dependent error components. The joint PDF of (W, V) is expressed by

$$f(w, v) = f_W(w) \cdot f_V(v) \cdot c(F_W(w), F_V(v)), \tag{10}$$

where $f_W(w)$ and $f_V(v)$ denote the PDFs of W and V, and $c(\cdot, \cdot)$ is the density of the copula.

The likelihood function of the copula-based SFM is written as:

$$L(\beta, \sigma_v, \sigma_w; Y, X) = \prod_{i=1}^{n} f_\varepsilon(\varepsilon) = \prod_{i=1}^{n} f_\varepsilon(logy_i - \beta x_i), \tag{11}$$

where $f_\varepsilon(\varepsilon)$ is the density function of ε, which can be obtained by the following steps [19]:

First, the joint density $f(w, \varepsilon)$ can be obtained using Equations (3) and (7). Transforming (w, v) to (w, ε), we can get the density function of (w, ε) [16]:

$$f(w, \varepsilon) = f_W(w) \cdot f_V(\varepsilon \mp w) \cdot c(F_W(w), F_V(\varepsilon \mp w)). \tag{12}$$

Second, marginalizing out W, we can obtain the density function of ε as

$$f_\varepsilon(\varepsilon) = \int_0^{+\infty} f(w, \varepsilon)dw, \tag{13}$$

or

$$f_\varepsilon(\varepsilon) = \mathrm{E}_W[f_V(\varepsilon \mp w) \cdot c(F_W(w), F_V(\varepsilon \mp w))], \tag{14}$$

where $\mathrm{E}_W[\cdot]$ represents the expectation function with respect to W.

Then, $f_\varepsilon(\varepsilon)$ can be approximated by

$$f_\varepsilon(\varepsilon) \approx \frac{1}{M} \sum_{m=1}^{M} f_V(\varepsilon \mp w_m) \cdot c(F_W(w_m), F_V(\varepsilon \mp w_m)), \tag{15}$$

where w_m, $m = 1, \ldots, M$ is a sequence of M random draws from a particular distribution, such as a standard half-normal distribution. The technology efficiency TE could be derived as

$$\begin{aligned} TE &= \mathrm{E}_W[\exp(-W)|\varepsilon] = \frac{1}{f_\varepsilon(\varepsilon)} \int_0^{+\infty} \exp(-W) f(w, \varepsilon) dw \\ &= \frac{\mathrm{E}_W[\exp(-W) f_V(\varepsilon \mp w) c(F_W(w), F_V(\varepsilon \mp w))]}{\mathrm{E}_W[f_V(\varepsilon \mp w) c(F_W(w), F_V(\varepsilon \mp w))]}. \end{aligned} \tag{16}$$

2.3. Simultaneous Stochastic Frontier Model with Dependent Error Components

In simultaneous SFM, which has also been referred to as seemingly unrelated stochastic frontier regressions, the stochastic frontier functions are estimated simultaneously. In this subsection, we propose a copula-based simultaneous SFM which allows for dependence between the random noise and inefficiency components of each stochastic frontier function, as well as dependence between the composite errors of two equations. We named our proposed model as simultaneous SFM with dependent error components for short, in order to distinguish from the copula-based simultaneous SFM which was first introduced by Lai and Huang [9] and further developed by Huang et al. [13,14] (hereafter referred to as simultaneous SFM with dependent composite errors). The basic form of the simultaneous SFM with dependent error components is expressed by

$$logY_1 = X\beta + \varepsilon_1, \tag{17}$$

$$logY_2 = Z\alpha + \varepsilon_2, \tag{18}$$

where

$$\varepsilon_1 = V_1 \pm W_1, \tag{19}$$

$$\varepsilon_2 = V_2 \pm W_2, \tag{20}$$

the dependence of error terms could be modeled by three copula functions, such that

$$H_1(v_1, w_1) = C_1\big(F_{V_1}(v_1), F_{W_1}(w_1); \theta_1\big), \tag{21}$$

$$H_2(v_2, w_2) = C_2\big(F_{V_2}(v_2), F_{W_2}(w_2); \theta_2\big), \tag{22}$$

$$H_{12}(\varepsilon_1, \varepsilon_2) = C_{12}(F_{\varepsilon_1}(\varepsilon_{1i}), F_{\varepsilon_2}(\varepsilon_{2i}); \theta_{12}). \tag{23}$$

Similar to the copula-based SFM, the noise term $V_j \sim \mathcal{N}\big(0, \sigma_{V_j}^2\big)$, $j = 1, 2$, is usually assumed to obey normal distribution, while the distribution of the inefficiency term W_j, $j = 1, 2$, could follow a half-normal, exponential or gamma distribution. $F_\varepsilon(\cdot)$, $F_V(\cdot)$, and $F_W(\cdot)$ are the CDFs of ε, V, and W, respectively. The parameters θ_1, θ_2, and θ_{12} represent the parameters of the three copulas $C_1(\cdot, \cdot)$, $C_2(\cdot, \cdot)$, and $C_{12}(\cdot, \cdot)$, which model the dependences between V_1 and W_1, V_2 and W_2, and ε_1 and ε_2, respectively.

The simultaneous SFM with dependent composite errors, proposed by Lai and Huang [9], could be regarded as a special case of the simultaneous SFM with dependent error components when the error components are assumed to be independent, such that

$$C_1\big(F_{V_1}(v_1), F_{W_1}(w_1); \theta_1\big) = F_{V_1}(v_1) \times F_{W_1}(w_1), \tag{24}$$

$$C_2\big(F_{V_2}(v_2), F_{W_2}(w_2); \theta_2\big) = F_{V_2}(v_2) \times F_{W_2}(w_2). \tag{25}$$

The likelihood function of the simultaneous SFM with dependent error components can be written as:

$$L(\Omega; Y_1, Y_2, X, Z) = \prod_{i=1}^{N} c_{12}(F_{\varepsilon_1}(\varepsilon_{1i}), F_{\varepsilon_2}(\varepsilon_{2i})) f_{\varepsilon_1}(\varepsilon_{1i}) f_{\varepsilon_2}(\varepsilon_{2i}), \tag{26}$$

with

$$f_{\varepsilon_j}(\varepsilon_j) \approx \frac{1}{M} \sum_{m=1}^{M} f_{V_j}(\varepsilon_j \mp w_{jm}) c(F_{W_j}(w_{jm}), F_{V_j}(\varepsilon_j \mp w_{jm})), \tag{27}$$

where Ω denotes the total possible parameter space and w_{jm}, $j = 1, 2$ and $m = 1, \ldots, M$, is a sequence of M random draws from a specific distribution. Therefore, the log-likelihood function of the simultaneous SFM with dependent error components can be expressed by

$$\begin{aligned} logL(\Omega) &= \sum_{i=1}^{N} log[c_{12}(F_{\varepsilon_1}(\varepsilon_{1i}), F_{\varepsilon_2}(\varepsilon_{2i}))] + \sum_{i=1}^{N} log(f_{\varepsilon_1}(\varepsilon_{1i})) + \sum_{i=1}^{N} log(f_{\varepsilon_2}(\varepsilon_{2i})) \\ &= \sum_{i=1}^{N} log[c_{12}(F_{\varepsilon_1}(\varepsilon_{1i}), F_{\varepsilon_2}(\varepsilon_{2i}))] \\ &\quad + \sum_{i=1}^{N} log\left[\frac{1}{M} \sum_{m=1}^{M} f_{V_1}(\varepsilon_{1i} \mp w_{1im}) \cdot c_1\left(F_{W_1}(w_{1im}), F_{V_1}(\varepsilon_{1i} \mp w_{1im}) \right) \right] \\ &\quad + \sum_{i=1}^{N} log\left[\frac{1}{M} \sum_{m=1}^{M} f_{V_2}(\varepsilon_{2i} \mp w_{2im}) \cdot c_2\left(F_{W_2}(w_{2im}), F_{V_2}(\varepsilon_{2i} \mp w_{2im}) \right) \right]. \end{aligned} \tag{28}$$

3. Simulation Study

The major advantage of simulation studies is that they are helpful to evaluate the behavior of statistical models, as some "truth" is known from the data generating process. This helps us to compare the performance and quality of one model against its competing methods [45]. To check the reasonability of our proposed simultaneous SFM with dependent error components, we perform two simulation experiments in this section. In the first simulation experiment, we compare the performance of the copula-based SFM with conventional SFM under the "truth" that the error components are correlated and the inefficiency terms are known. The second simulation is performed to make a comparison between the performance of our proposed simultaneous SFM with dependent error components and the simultaneous SFM with dependent composite errors of Huang et al. [13,14].

3.1. Comparative Study of Copula-Based SFM and Conventional SFM

We conducted a simulation experiment to make a comparison between the performance of copula-based SFM introduced in Section 2.2 and conventional SFM, under the assumption that the error components were dependent. Our simulation was based on a simple SFM with single explanatory variable, expressed by

$$y = x'\beta + v + w, \tag{29}$$

where y is the output vector, x is the vector of a single explanatory variable, and β is the unknown parameter to be estimated. The composite error is expressed as $\varepsilon = v + w$, where the error component v represents the statistical noise, while w stands for the inefficiency term. Here the marginal distribution of w was assumed to be a half-normal distribution with $\left| \mathcal{N}(0, \sigma_w^2) \right|$, while the random noise term v was assumed to be normal distributed $\mathcal{N}(0, \sigma_v^2)$. We assumed v and w to be dependent in the true model.

The data-generating mechanism consisted of the following steps:

- Set up the values for parameters. To generate a simulated data set, the true parameters of the SFM were fixed as $\beta = 10$, $\sigma_w = 0.7$, and $\sigma_v = 0.7$. We chose the Gaussian copula to obtain the correlation between v and w with the copula parameter set to be $\rho = 0.7$.

- Simulate distributions of w and v. We first simulated the distribution of w (F_w) by generating a sequence of 1000 random draws from the Halton sequence. Then, the conditional distribution of v given w (F_v) was simulated by a Gaussian copula $C_{gau}(v|w)$ using the "BiCopCondSim" function in the R software.

- Obtain simulated data of w and v from their simulated distributions. The inefficiency term w was generated by computing the inverse of the half-normal distribution with $\left|\mathcal{N}\left(0, \sigma_w^2\right)\right|$, given the distribution F_w obtained in the last step; the statistical noise term v was computed as the inverse of the normal distribution with $\mathcal{N}\left(0, \sigma_v^2\right)$, given the distribution F_v. The composite errors were then computed by $\varepsilon = v + w$.

- Simulate data of variables x and y. The data of the explained variable x was generated from uniform random numbers on the interval [0, 1], while the dependent variable y was generated according to Equation (29).

We generated 500 data sets of size $n = 200$, based on the above process. We estimated the conventional SFM by the package "frontier" in the R software and used the estimated coefficients as the starting values of β, σ_w, and σ_v to estimate the copula-based SFM and conventional SFM. Next, we estimated the two models 500 times each using maximum simulated likelihood, see details in Greene [46]. The simulated log-likelihood was maximized using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm in R software. The performance of two models was then compared by the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

The true values (True) of parameters (Para) and the summary statistics of estimated parameters by simulation are summarized in Table 1. The estimation accuracy of parameters from the two models were compared by the values of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The overall MAE and MAPE for parameters β, σ_w, σ_v, and ρ from the copula-based SFM is 0.057 and 0.05, which is lower than the overall MAE (0.316) and MAPE (0.428) from the conventional SFM. In particular, the MAE and MAPE of σ_w estimated from copula-based SFM were obviously lower than the conventional SFM, which indicates that the parameters σ_w estimated by copula-based SFM were much more accurate than conventional SFM. Furthermore, the average AIC and BIC of copula-based SFM were both lower than the conventional SFM, implying that copula-based SFM outperformed the conventional SFM and gave a better fit to data.

Table 1. Simulation results of the conventional SFM and copula-based SFM.

Para	True	Mean	Max	Min	Median	95% CI	MAE	MAPE
				Conventional SFM				
β	10.0	9.981	10.154	9.794	9.983	[9.859, 10.107]	0.054	0.005
σ_w	0.7	1.566	3.587	0.003	1.563	[1.118, 2.074]	0.870	1.243
σ_v	0.7	0.695	0.790	0.602	0.695	[0.638, 0.758]	0.025	0.036
AIC		565.1	616.5	507.8	565.2	[530.8, 600.0]		
BIC		575.0	626.4	517.7	575.1	[540.7, 609.9]		
Overall							0.316	0.428
				Copula-based SFM				
β	10.0	9.993	10.420	9.556	9.996	[9.751, 10.218]	0.093	0.009
σ_w	0.7	0.698	0.997	0.413	0.703	[0.530, 0.859]	0.068	0.097
σ_v	0.7	0.700	0.822	0.577	0.699	[0.621, 0.781]	0.031	0.045
ρ	0.7	0.700	0.804	0.518	0.703	[0.608, 0.772]	0.034	0.049
AIC		432.4	494.5	354.0	434.0	[391.9, 470.6]		
BIC		445.6	507.6	367.2	447.2	[405.1, 483.8]		
Overall							0.057	0.050

Note: 95% CI: 95% confidence intervals; MAE: Mean Absolute Error; MAPE: Mean Absolute Percentage Error; AIC: Akaike Information Criterion.

From the simulation study, the estimated parameters β, σ_w, σ_v, and ρ from the copula-based SFM were found to be closer to their true values, with mean values of 9.993, 0.698, 0.700, and 0.700, respectively. It is remarkable that the mean (0.698) and median (0.703) of σ_w from the copula-based SFM were much closer to the true value (0.7), compared with the mean (1.566) and median (1.563) of σ_w from the conventional SFM. Moreover, the parameter σ_v estimated by copula-based SFM was also more precise than by conventional SFM. Therefore, ignoring dependence between the error components w and v in conventional SFM may lead to biased estimations, which was also established by Smith [16], Sriboonchitta et al. [41] and Wiboonpongse et al. [18].

Further, Figure 1a,b plot the histogram and kernel density of the estimated parameters by conventional SFM and copula-based SFM, respectively. Generally speaking, the kernel density of each parameter from copula-based SFM fit well to the histograms and was quite close to a normal distribution. However, the estimated parameter σ_w by conventional SFM obviously deviated from its true value, as shown in Figure 1a. On the contrary, the σ_w estimated by copula-based SFM was much closer to its true values, as shown in Figure 1b. Therefore, the copula-based SFM outperformed the conventional SFM when the error components were correlated.
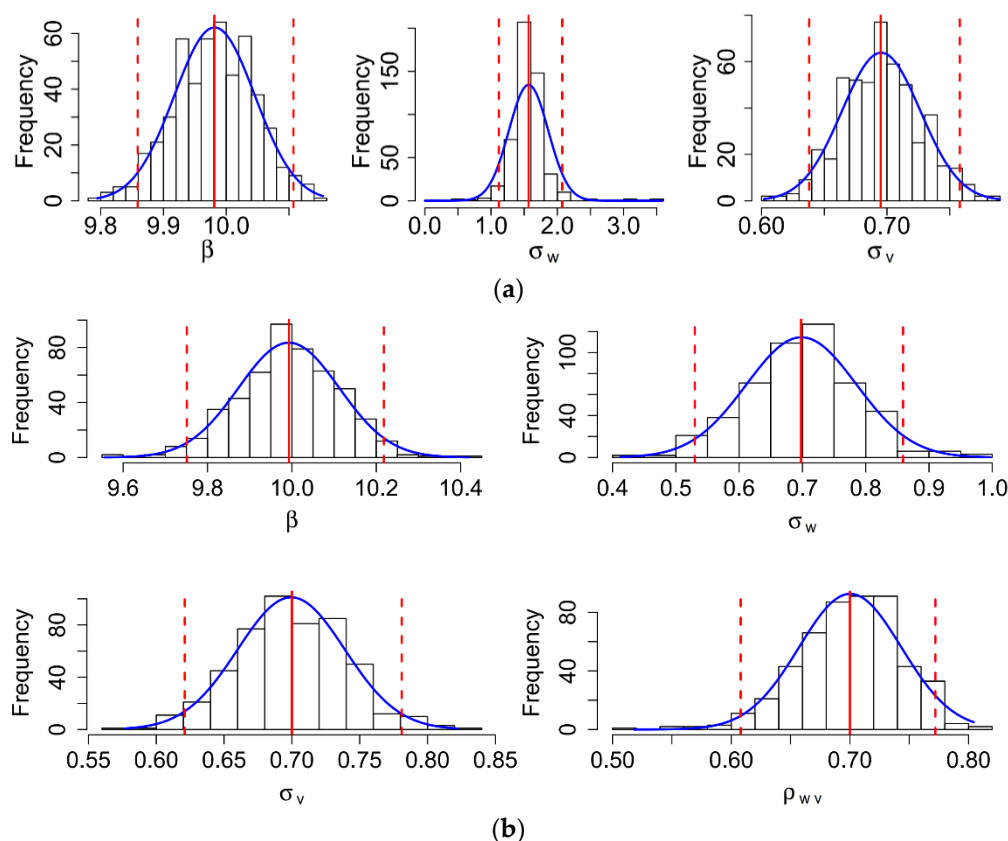

Figure 1. (**a**) Histogram and kernel density of parameters β, σ_w, and σ_v from conventional SFM. (**b**) Histogram and kernel density of parameters β, σ_w, σ_v, and ρ_{wv} from copula-based SFM.

3.2. Comparative Study of Two Copula-Based Simultaneous SFMs

Our second simulation study was aimed at comparing the performance of two copula-based simultaneous SFMs: simultaneous SFM with dependent composite errors and simultaneous SFM with dependent error components. The simultaneous SFM in this simulation study could be expressed by

$$y_1 = \alpha_0 + x_1'\alpha_1 + v_1 + w_1, \tag{30}$$

and

$$y_2 = \beta_0 + x_2'\beta_1 + v_2 + w_2, \tag{31}$$

where $\varepsilon_1 = v_1 + w_1$ and $\varepsilon_2 = v_2 + w_2$.

We allowed for dependence between the error components, such that v_1 was correlated with w_1 and v_2 was correlated with w_2. Meanwhile, the composite errors ε_1 and ε_2 were also assumed to be dependent. The marginal distribution of the inefficiency terms w_1 and w_2 were assumed to be half-normally distributed ($\left| \mathcal{N}\left(0, \sigma_{wj}^2\right) \right|, j = 1, 2$), while the noise terms v_1 and v_2 were assumed to be normally distributed ($\mathcal{N}\left(0, \sigma_{vj}^2\right), j = 1, 2$).

The data generation process can be summarized as follows:

- Simulate values of the parameters. The coefficients $\alpha_0, \alpha_1, \beta_0$, and β_1 were generated from uniformly distributed random numbers on the interval $[5, 10]$, the standard deviations of inefficiency terms σ_{w1} and σ_{w2} were generated from uniform random numbers on the interval $[2, 3]$, and the standard deviations of noise terms σ_{v1} and σ_{v2} were simulated from uniform random numbers on the interval $[0.5, 1]$. The dependence between v_1 and w_1, v_2 and w_2, and ε_{1i} and ε_{2i} were modeled by Gaussian copulas, where the copula parameters ρ_1, ρ_2, and ρ_{12} were simulated from uniform random numbers on the interval $[0.7, 0.95]$.
- Simulate distributions of w_1 and v_1, and w_2 and v_2. To simulate the data of w_1 and v_1, we first simulated the distribution of w_1 (F_{w1}) by generating a sequence of 500 random draws from the Halton sequence. Next, we simulated the conditional distribution of v_1 given w_1 (F_{v1}) from a Gaussian copula $C_{gau}(v_1|w_1)$ using the "BiCopCondSim" function in R software, setting up the copula parameter to be ρ_1. The distributions of w_2 (F_{w2}) and v_2 (F_{v2}) were simulated following the same procedure as w_1 and v_1, where the copula parameter was set as ρ_2.
- Generate values for $w_1, v_1, \varepsilon_1, w_2, v_2$, and ε_2. In this step, we generated the values of w_1 and v_1, as well as w_2 and v_2, given their distributions simulated in the last step. The inefficiency term w_1 (w_2) was computed by the inverse of the half-normal distribution with a mean of zero and standard deviation of σ_{w1} (σ_{w2}); the noise term v_1 (v_2) was computed by the inverse of the normal distribution with a mean of zero and standard deviation of σ_{v1} (σ_{v2}); and the composite errors were computed by $\varepsilon_1 = v_1 + w_1$ and $\varepsilon_2 = v_2 + w_2$.
- Generate values for variables x_1 and y_1, and x_2 and y_2. The explained variables x_1 and x_2 were simulated as uniform random numbers on the interval $[0, 1]$, while the dependent variables y_1 and y_2 were calculated according to Equations (30) and (31), respectively.

Following the above steps, we generated 200 data sets of size $n = 500$ for each model. We first estimated the conventional SFM by the package "frontier" in R software, and used the estimated coefficients as the starting values of $\alpha_0, \alpha_1, \sigma_{w1}, \sigma_{v1}, \beta_0, \beta_1, \sigma_{w2}$, and σ_{v2}; the starting values of three copula parameters ρ_1, ρ_2 and ρ_3 were all set at 0.5. Then, we estimated the simultaneous SFM with dependent error components and simultaneous SFM with dependent composite errors using maximum simulated likelihood and compared the quality of the two models by AIC and BIC. The simultaneous SFM with dependent error components in this simulation study was supposed to be the true model, as the data were generated based on the assumption that the error components were correlated. Thus, the simultaneous SFM with dependent composite errors was mis-specified.

In Table 2, we summarize the statistics of AIC and BIC calculated from the two copula-based simultaneous SFMs. The average AIC and BIC of the simultaneous SFM with dependent error components are both lower than the simultaneous SFM with dependent composite errors, with an average difference of −228.2 (−219.7) for AIC (BIC) between the two models. Further, the difference of AICs between the two models is plotted in Figure 2. During 200 simulations, the AICs of simultaneous SFM with dependent error components were smaller than the AICs of simultaneous SFM with dependent composite errors for 197 times. We can conclude that the mis-specified model (with

larger AIC and BIC) lost more information and, thus, had lower qualities compared to the true model. Therefore, if the error components of SFM are correlated, the simultaneous SFM with dependent error components provided a better fit for the data.

Table 2. Summary statistics of AIC and BIC of two copula-based simultaneous SFM.

Criteria	Mean	Max	Min	Median	95% CI
Simultaneous SFM with dependent error components					
AIC	3031.3	3534.4	2273.7	3035.5	[2491.4, 3504.1]
BIC	3077.7	3580.8	2320.1	3081.9	[2537.8, 3550.4]
Simultaneous SFM with dependent composite errors					
AIC	3259.5	4170.1	2279.5	3275.6	[2520.8, 4046.0]
BIC	3297.4	4208.0	2317.4	3313.5	[2558.7, 4083.9]
D.AIC	−228.2	0.662	−688.8	−207.8	[−636.0, −1.405]
D.BIC	−219.7	9.091	−680.4	−199.4	[−627.6, 7.024]

Note: 95% CI: 95% confidence intervals; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; D.AIC (D.BIC): AICs (BICs) of simultaneous SFM with dependent error components minus AICs (BICs) of simultaneous SFM with dependent composite errors.
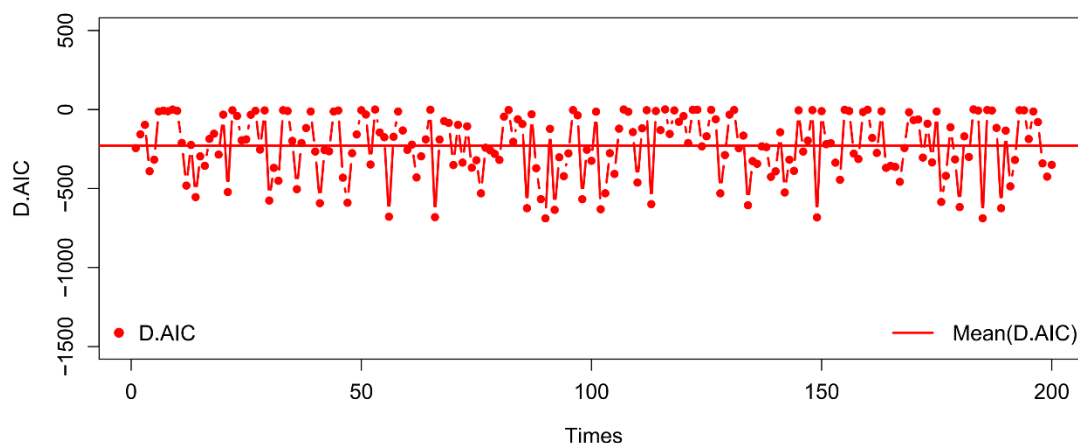

Figure 2. Differences between AICs of two copula-based simultaneous SFMs. Note: D.AIC: AICs of simultaneous SFM with dependent error components minus AICs of simultaneous SFM with dependent composite errors.

3.3. Brief Summary

The results of the two simulation studies confirmed reasonability and feasibility of allowing for possible correlations between the error components in SFM as well as simultaneous SFM. Conclusions from the first simulation study challenged the ubiquitous assumption of independence between statistical noise and inefficiency components in conventional SFM. When the two error components were correlated, estimation by conventional SFM led to relatively larger errors as compared with the copula-based SFM which modelled dependence of the error components by copula functions. Furthermore, for 200 simulations in the second simulation study, the AICs of simultaneous SFM with dependent error components were always lower than the AICs of simultaneous SFM with dependent composite errors. This result strongly supported that the simultaneous SFM with dependent error components could provide a better fit to the data as compared with the simultaneous SFM with dependent composite errors developed by Huang et al. [13,14]. Therefore, it is advisable to relax the restrictive assumption of independence between two error components when analyzing real economic problems using either SFM and/or simultaneous SFM.

4. An Application to the Chinese Banking Industry

In this section, the simultaneous SFM with dependent error components (as proposed in Section 2) was applied to a balanced panel dataset of the Chinese banking industry. We estimated market power, cost efficiency and meta-frontier technology gap ratio of 37 Chinese commercial banks covering the period 2013 to 2018. Meanwhile, estimation based on simultaneous SFM with composite errors and the single-equation SFM was also carried out for comparison.

4.1. Model and Data

In this application, we jointly estimated a cost frontier function and an output price frontier function using the simultaneous SFM with dependent error components, in which the total costs and output price were considered as dependent variables. The simultaneous SFM with dependent error components combined a translog cost function and an output price function using copulas, which can be specified as

$$
\begin{aligned}
lnTC = {} & \alpha_0 + \alpha_1 lnY_{it} + \tfrac{1}{2}\alpha_2 (lnY_{it})^2 + \sum_{k=1}^{3} \lambda_k lnP_{k,it} + \tfrac{1}{2}\sum_{k=1}^{3}\sum_{h=1}^{3} \beta_{kh} lnP_{k,it} lnP_{h,it} \\
& + \sum_{k=1}^{3} \gamma_k lnY_{it} lnP_{k,it} + \omega_1 Trend + \tfrac{1}{2}\omega_2 Trend^2 + \omega_3 Trend \times lnY_{it} \\
& + \sum_{k=1}^{3} \varphi_k Trend \times lnP_{k,it} + \varepsilon_{1it},
\end{aligned}
\tag{32}
$$

and

$$
P_{it}^* = MC_{it} + \varepsilon_{2it},
\tag{33}
$$

where $\varepsilon_{1it} = v_{1it} + w_{1it}$, $\varepsilon_{2it} = v_{2it} + w_{2it}$, and

$$
F(v_{1it}, w_{1it}) = C\big(F_{V_1}(v_{1it}), F_{W_1}(w_{1it})\big),
\tag{34}
$$

$$
F(v_{2it}, w_{2it}) = C\big(F_{V_2}(v_{2it}), F_{W_2}(w_{2it})\big),
\tag{35}
$$

$$
F(\varepsilon_{1it}, \varepsilon_{2it}) = C(F_{\varepsilon_1}(\varepsilon_{1it}), F_{\varepsilon_2}(\varepsilon_{2it})).
\tag{36}
$$

Equation (32) represents the translog cost function of bank i at time t (which is also referred to as the cost frontier function), where $lnTC$ donates the total cost; Y represents for the output; P_k ($k = 1, 2, 3$) implies the input prices of labor, capital, and funds, respectively; and $trend$ is the time trend, indicating technical changes over time. The error component v_{1it} represents the random noise, while w_{1it} stands for the non-negative inefficiency term. Equation (33) is the so-called output price frontier function, where P_{it}^* and MC_{it} stand for the output price and the marginal cost of bank i at time t, respectively [14,25]. The error component v_{2it} stands for the random noise in the price frontier, while w_{2it} is a non-negative random variable measuring the extent to which price deviates from MC_{it}.

The implied marginal cost MC_{it} is calculated by taking the partial derivative of total costs TC_{it} with respect to output Y_{it}:

$$
MC = \frac{\partial TC_{it}}{\partial Y_{it}} = \frac{TC_{it}}{Y_{it}}\left[\alpha_1 + \alpha_2 lnY_{it} + \sum_{k=1}^{3} \gamma_k lnP_{k,\,it} + \omega_3 Trend\right].
\tag{37}
$$

Thus, it is crucial to allow the total cost (Equation (32)) and output price (Equation (33)) to be correlated. In the simultaneous SFM with dependent error components, the dependence between the error components w_{1it} and v_{1it}, and w_{2it} and v_{2it} were modelled by copulas, as shown in Equations (34) and (35). Moreover, the composed error terms ε_{1it} and ε_{2it} were also permitted to be dependent, following Equation (36).

Based on the estimation results of the simultaneous SFM with dependent error components, a set of indices to measure market power, cost efficiency, and the meta-frontier technology gap ratio of banks could be derived. First, the Lerner Index (LI), a well-established measure of market power (competition) of firms, was calculated as $E(u_{2t}|\varepsilon_{2it})/P_{it}$. Second, the measure of scale economies (SC) could be obtained by $\partial lnTC/\partial lnY$, which is the term in the parenthesis in Equation (37). Further, the cost efficiency (CE), technology gap ratio (TGR) as well as meta-frontier cost efficiency (MCE) could be obtained; for more details, see Huang et al. [13,14].

We used a balanced panel dataset comprised of 222 observations for 37 Chinese commercial banks from 2013 to 2018, to avoid selection bias from unbalanced data or non-random sample. Our sample included all the 6 state-owned banks and 12 joint-stock banks of China. Furthermore, we selected 19 Chinese city commercial banks and rural commercial banks which were listed on A-share or H-share stock market by the end of 2018. The total assets of the sample banks account for approximately 68% of the total assets of the Chinese banking industry during the sample period, which indicates that these 37 banks are representative of the Chinese banking industry. The data of all variables were gathered from the Bankscope database. The missing values were manually recorded from annual reports of individual banks. We divided the 37 Chinese banks into four sub-samples: (a) state-owned commercial banks (SOCBs), (b) joint-stock commercial banks (JSCBs), (c) city commercial banks (CCBs) and (d) rural commercial banks (RCBs), according to their ownership types. The output of banks was proxied by their total assets (Y). The price of labor (P_1), price of capital (P_2), and the price of funds (P_3) were calculated as the ratio of personnel expenses to number of employees, the ratio of operating expenses to total fixed assets, and the ratio of interest expenses to all types of deposits, respectively. Total cost (TC) was calculated as the sum of personnel expenses, operating expenses, and interest expenses. The output price P^* was defined as the ratio of total revenue (TR) to output (Y). Following Shamshur and Weill [47] and Huang et al. [13,14], we used the consumer price index (CPI) of China to deflate all variables from nominal values into real values with base year 2013 = 100. The descriptive statistics of variables are summarized in Table 3.

Table 3. Descriptive statistics of used variables.

Variables	Entire Sample	SOCBs	JSCBs	CCBs	RCBs
Total revenues (TR)	1.7778×10^8 (2.5983×10^8)	6.7775×10^8 (2.6755×10^8)	1.9337×10^8 (9.0484×10^7)	3.2843×10^7 (2.5205×10^7)	1.4155×10^7 (1.3939×10^7)
Total assets (Y)	3.8862×10^9 (6.0574×10^9)	1.5698×10^{10} (6.5007×10^{10})	3.7259×10^9 (1.8339×10^9)	7.0699×10^8 (5.6708×10^8)	2.8958×10^8 (2.9301×10^8)
Total costs (TC)	1.0856×10^8 (1.4852×10^8)	4.0271×10^8 (1.2239×10^8)	1.2199×10^8 (5.2576×10^7)	2.1694×10^7 (1.6508×10^7)	9.1620×10^6 (8.9081×10^6)
Price of labor (P_1)	345.269 (84.186)	246.245 (38.131)	397.954 (73.877)	347.937 (77.403)	349.817 (63.075)
Price of capital (P_2)	1.049 (0.743)	0.807 (0.738)	1.254 (0.597)	1.083 (0.875)	0.865 (0.477)
Price of funds (P_3)	0.027 (0.006)	0.019 (0.004)	0.028 (0.005)	0.029 (0.005)	0.0262 (0.005)
Output price (P^*)	0.049 (0.006)	0.044 (0.005)	0.053 (0.005)	0.048 (0.005)	0.050 (0.006)
No. of bank	37	6	10	15	6
Obs.	222	36	60	90	36

Note: Sample means are reported (Thousands of real Chinese Yuan with base year 2013); Numbers in parentheses are standard deviations.

It was observed that SOCBs played the dominant role in the Chinese banking industry. Regarding bank size, SOCBs were largest among all types of banks and were roughly four times as large as JSCBs

on average in terms of total assets. On the contrary, the sizes of CCBs and RCBs were much smaller compared with SOCBs and JSCBs. However, the three input prices and output price of SOCBs were lowest as compared with JSCBs, CCBs and RCBs during the sample period.

4.2. Estimation Results

In our application, the random noise terms v_{1it} and v_{2it} were assumed to be normally distributed ($\mathcal{N}\left(0,\ \sigma_{vi}^2\right), i = 1,\ 2$), while the inefficiency terms w_{1it} and w_{2it} were assumed to be either half-normally (HN) or exponentially (EX) distributed. We considered Gaussian (G) and Frank (F) copulas to model dependence between random error and inefficiency terms. The simultaneous SFM with dependent error components was estimated by maximum simulated likelihood using 500 draws from the Halton sequence. The best-fitting model was selected according to the AIC values.

Figure 3 plots the AIC values of all considered models. As shown, the best-fitting model (GFG) is the one based on a Gaussian copula to capture dependence between v_1 and w_1, a Frank copula for dependence between v_2 and w_2, and a Gaussian copula for dependence between ε_1 and ε_2, with the lowest AIC value (-2255.7) among all considered models. It is worth mentioning that the models in which the inefficiency terms w_1 and w_2 were described by a HN distribution were all superior to the models using an EX distribution, indicating that the marginal distribution of inefficiency terms has a crucial influence on the goodness of model fit, especially in our case.
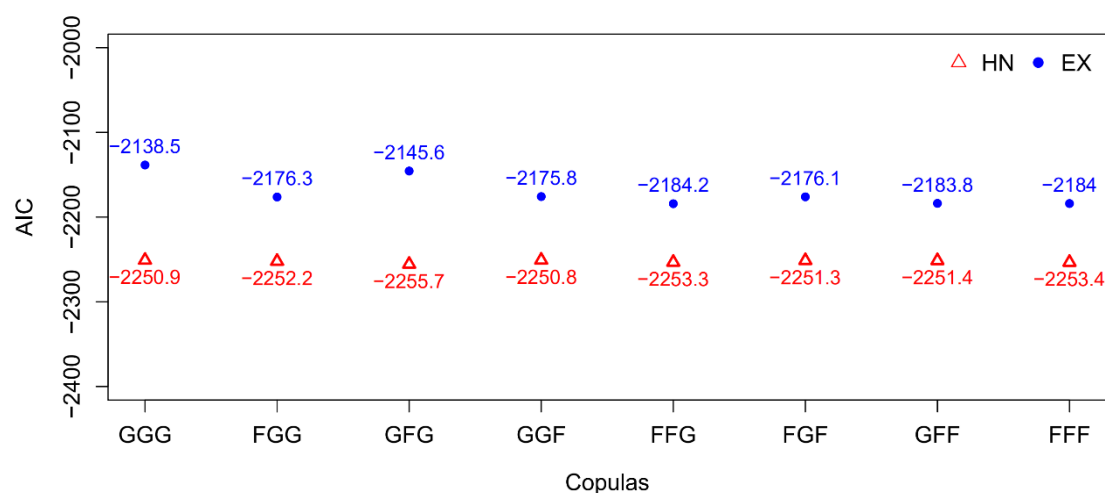

Figure 3. AIC values for each copula-based simultaneous SFM (the symbol "G" stands for "Gaussian copula"; "F" represents "Frank copula"; and "HN" and "EX" represent half-normal and exponential distributions, respectively. The three copulas are listed, according to the order of the dependence between v_1 and w_1, v_2 and w_2, and ε_1 and ε_2).

The estimation results of the cost function by the best-fitting model of the simultaneous SFM with dependent error components (SSFMDEC) are shown in Table 4. Meanwhile, we provide the result obtained by the simultaneous SFM with dependent composite errors (SSFMDCE) of Huang et al. [13,14] where the dependence between v_1 and w_1, and v_2 and w_2 were both modelled by independent copulas. To draw a solid comparison with our simultaneous SFM with dependent error components, the dependence between ε_1 and ε_2 in the simultaneous SFM with dependent composite errors was also modelled by a Gaussian copula and the inefficiency terms were assumed to have HN distributions. The estimation results of the conventional single-equation SFM is also provided for comparison.

The results immediately yield some clear conclusions: First, most of the estimated coefficients of the cost function by SSFMDEC and SSFMDCE were significant at 1% level of significance, while many parameters failed to be significant in the conventional single-equation SFM. Thus, the two copula-based simultaneous SFMs (SSFMDEC and SSFMDCE) provided more efficient results as compared to the single-equation SFM. Second, the AIC of the single-equation SFM had a much larger value than

the other two models, indicating that both versions of simultaneous SFM were preferable to the single-equation SFM. Third, the AIC of SSFMDEC was lower than the AIC of SSFMDCE. Thus, the simultaneous SFM with dependent error components outperformed the simultaneous SFM with dependent composite errors. Finally, it is important to note that the Gaussian copula parameter ρ_1 and Frank copula parameter θ_2 in SSFMDEC were both significant, which indicate that it is rational to allow dependence between the error components in simultaneous SFM.

Table 4. Parameter estimation results of the cost function.

Models	SSFMDEC		SSFMDCE		Single-Equation SFM	
Variables	Parameter	S.E.	Parameter	S.E.	Parameter	S.E.
Constant	13.123 ***	0.0050	9.8454 ***	0.0060	11.707 **	5.1775
lnY	0.2512 ***	0.0002	0.3806 ***	0.0003	0.1712	0.2022
$lnY \times lnY$	0.0344 ***	0.0000	0.0494 ***	0.0004	0.0522 ***	0.0049
lnP_2	0.5719 ***	0.0009	−0.5109 ***	0.0010	−0.5108	0.4778
lnP_3	1.8974 ***	0.0005	2.1181 ***	0.0006	2.0425 *	1.0721
$lnP_2 \times lnP_2$	0.1526 ***	0.0003	0.1205 ***	0.0003	0.1177 ***	0.0276
$lnP_3 \times lnP_3$	0.1005 ***	0.0001	0.2423 ***	0.0001	0.1981	0.1416
$lnP_2 \times lnP_3$	−0.0483 ***	0.0002	−0.1786 ***	0.0055	−0.1716 *	0.0881
$lnY \times lnP_2$	0.0057 ***	<0.0001	0.0220 ***	<0.0001	0.0224 **	0.0072
$lnY \times lnP_3$	−0.0101 ***	<0.0001	0.0264 ***	<0.0001	0.0109	0.0218
$trend$	−0.0921 ***	0.0184	−0.2650 ***	0.0016	−0.2707	0.1729
$trend \times trend$	0.0102 **	0.0036	0.0081 ***	0.0012	0.0083	0.0052
$trend \times lnY$	−0.0035 ***	<0.0001	−0.0002 ***	<0.0001	−0.0017	0.0026
$trend \times lnP_2$	0.0219 ***	0.0002	0.0092 ***	0.0028	0.0106	0.0080
$trend \times lnP_3$	−0.0254 ***	0.0001	−0.0294 ***	0.0005	−0.0340 *	0.0174
σ					0.0110 **	0.0043
γ					0.4587	0.4080
σ_{w1}	0.0231	0.0159	0.0915 ***	0.0242		
σ_{w2}	0.0768 ***	0.0068	0.0716 ***	0.0103		
σ_{v1}	0.0219 ***	0.0004	0.0162 ***	0.0008		
σ_{v2}	0.0136 ***	0.0008	0.0012	0.0008		
ρ_1	0.9800 ***	<0.0001				
θ_2	−28.346 ***	2.3940				
ρ_{12}	0.0148	0.0736	0.1297	0.0708		
AIC	−2255.67		−1907.41		−413.99	

Note: Significance at the 0.01, 0.05, and 0.10 levels are indicated by ***, **, *, respectively; SSFMDEC is short for simultaneous SFM with dependent error components, SSFMDCE is short for simultaneous SFM with dependent composed errors; P_1 is arbitrarily selected as the numeraire to satisfy the homogeneity restriction in input prices; In single-equation SFM, $\sigma = \sqrt{\sigma_w^2 + \sigma_v^2}$ and $\gamma = \sigma_w^2/\sigma^2$.

4.3. Various Measures of Interests

The parameters estimated by the simultaneous SFM with dependent error components were then applied to compute a set of indicators to analyze the level of competition, cost efficiency and technology gap ratio of Chinese banks, such as the Lerner Index (LI), scale economies (SC), cost efficiency (CE), technology gap ratio (TGR) as well as meta-frontier cost efficiency (MCE).

4.3.1. The Lerner Index (LI) and Scale Economies (SC)

Summary statistics of LI and SC are presented in Table 5. A bank's LI reflects relative makeup of the market output price (P^*) over marginal cost (MC), which is defined by $LI_{it} = \left(P^*_{it} - MC_{it}\right)/P^*_{it}$. LI is a measure of market power of firms and can be regarded as the inverse of market competition level. The values of LI range from 0 (perfect competition) to 1 (pure monopoly) [48]. We obtained an average LI of 33.5% for the entire sample, with group mean ranging from 31.1–38.0%. This result is close to the statistics of the Federal Reserve Economic Data (FRED), which reported an average LI of 34.9% in the banking market of China from 1997 to 2014. Compared with the banks of developed countries during

the same period (such as the average LI of 26.8% for banks in United States), Chinese banks possessed higher market power. Second, there were some discrepancies in competition levels between different types of banks. The average LI was 31.1% for SOCBs, 31.4% for JSCBs, 34.1% for CCBs, and 38.0% for RCBs. The market competition of both SOCBs and JSCBs were highest and did not differ substantially from each other, followed by CCBs; whereas RCBs had the lowest competition level on average. This finding indicates that competition of the Chinese banking industry mainly exists in SOCBs and JSCBs. The reason for this is that RCBs and CCBs are limited within a certain territorial area, typically serving local residents and small enterprises. Compared with SOCBs and JSCBs, CCBs and RCBs have the advantages of more flexible operation modes, lower transaction costs, and better understanding of local conditions. Hence, RCBs and CCBs were found to have greater market power than SOCBs and JSCBs. Our results are in line with the findings by Tan [49] and Fungáčová [21], who documented that the competition of SOCBs and JSCBs were higher than CCBs.

With respect to the economies of scale, a company operates under diminishing, constant, or increasing returns to scale if SC is greater than, equal to, or less than one, respectively [50]. The mean and median values of SC for the entire sample were greater than 1, indicating that Chinese banks operated under slightly decreasing returns to scale overall. Among them, all values of SC for SOCBs and JSCBs were greater than 1. Thus, diseconomies of scale occurred in SOCBs and JSCBs. On the contrary, CCBs and RCBs achieved overall economies of scale, also known as increasing returns to scale. According to Berger and Humphrey [51], small-sized banks can obtain scale economies by increasing their size, while further increases in size may result in diseconomies of scale after a certain point. This statement is further supported by the findings of Barros et al. [52] and Athanasoglou et al. [53]. Thus, our finding is acceptable because the average sizes of SOCBs and JSCBs were much larger than CCBs and RCBs based on the *TR* and *Y* values, as shown in Table 3.

Table 5. Summary statistics of Lerner Index and Scale Economies.

Groups	Entire Sample	SOCBs	JSCBs	CCBs	RCBs
		Lerner Index (LI)			
Mean	0.335	0.311	0.314	0.341	0.380
S.D.	0.067	0.090	0.060	0.053	0.060
Median	0.337	0.327	0.318	0.339	0.392
Min	0.104	0.104	0.154	0.197	0.258
Max	0.487	0.417	0.460	0.475	0.487
		Scale Economies (SC)			
Mean	1.019	1.103	1.053	0.989	0.953
S.D.	0.056	0.015	0.021	0.025	0.033
Median	1.007	1.110	1.060	0.989	0.935
Min	0.920	1.073	1.002	0.944	0.920
Max	1.121	1.121	1.077	1.038	1.003
Obs.	222	36	60	90	36

Next, we take a closer look at the dynamic characteristics of the market power of Chinese banks. Figure 4 plots the average LI for different types of banks by year. Overall, the increasing average LI provides evidence of enhanced market power of Chinese banks during the period 2013–2018, both as a whole or specific groups. Second, the market power of RCBs remained the highest among all types of banks each year, followed by CCBs; while the market power of SOCBs remained low and did not vary largely across the whole period. However, it is noticeable that the average LI of JSCBs showed some variation. The market power of JSCBs remained lowest before 2014, then increased from 2014 to 2016 and reached the average level of the whole sample. Third, it is worth noting that the market power of JSCBs, CCBs, and RCBs all decreased in 2016, indicating increased competition among these banks. This result may be due to the tax reform policy of China implemented in 2016 to replace business tax

(BT) by value-added tax (VAT) in several service sectors, including financial service sector. The tax burden of many banks and financial companies increased at the beginning of the tax reform thereby leading to an increase in operating costs for banks. Therefore, the market power of Chinese banks, especially for the smaller banks, decreased in 2016.
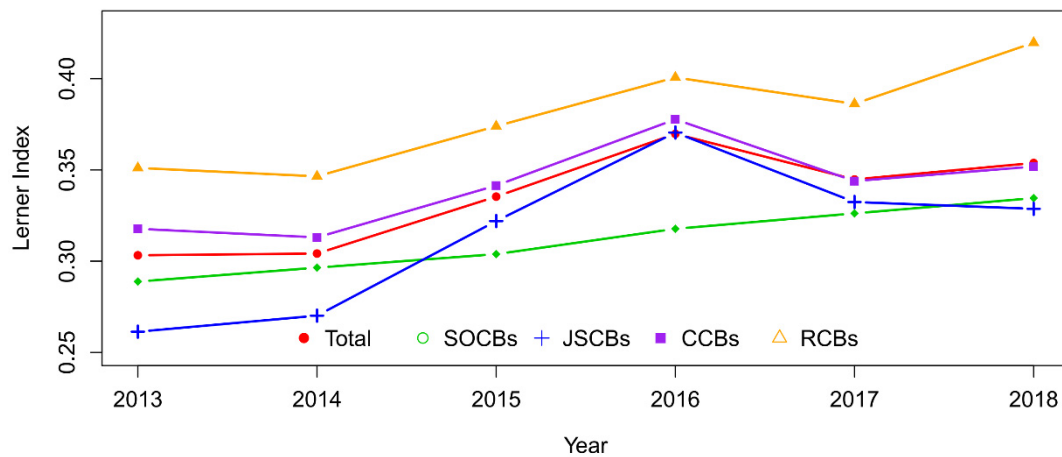

Figure 4. The average Lerner Index for different types of banks by year.

The gaps in market power amongst banks, as measured by the difference between maximum LI and minimum LI within each group, are plotted in Figure 5. The difference of LI tended to reduce for every type of bank, indicating that the gap of market power amongst banks has narrowed in recent years. Regarding SOCBs, the gap of market power remained relatively stable from 2013 to 2016. However, the gap has narrowed since 2016 due to the ever-increasing market power of the Postal Savings Bank of China (PSBC). The Industrial and Commercial Bank of China (ICBC) had the highest market power among all SOCBs from 2013 to 2017, whereas the market power of PSBC was lowest in those years. The gap of market power for JSCBs remained stable, with a slightly downward trend, implying that the market power of joint stock banks was changing. Obviously, the market power of the China Merchants Bank (CMBC) remained the highest every year, while the China Bohai Bank (CBHB) faced the highest level of competition for most years (except for 2017). Turning to CCBs and RCBs, the differences of LI for CCBs and RCBs shrank from 2013 to 2018, indicating that the gap of market power among banks within the two groups was diminishing. In addition, banks with the highest and lowest market power varied amongst different banks for CCBs and RCBs. SOCBs are the mainstay of China's commercial banks, with respect to asset and loan sizes, and most state revenues and expenses are handled by state-owned banks, while the operation of JSCBs is largely attributed to contributions of shareholders. Thus, the market power of SOCBs and JSCBs were relatively stable, leading to the fact that the gap of market power within the two types of banks varied only slightly over the period under consideration. On the contrary, CCBs and RCBs are mainly controlled by local governments and enterprises, with the characteristics of large quantities, regional restriction and instability. Therefore, there were more obvious changes of market power among CCBs and RCBs.
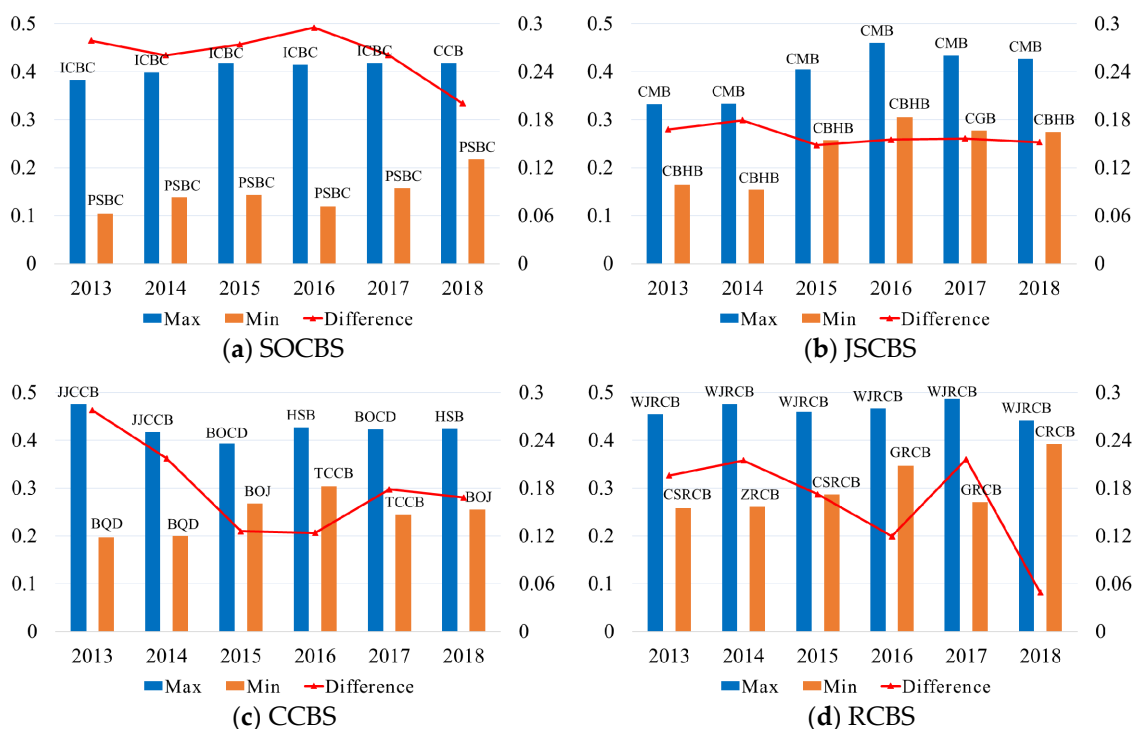
Figure 5. The values of maximum, minimum, and difference of Lerner Index for the four studied bank groups. (The left-hand ordinate indicates the LI values and the right-hand ordinate indicates the differences between maximum and minimum LI within each group; CCB: China Construction Bank, CGB: China Guangfa Bank, JJCCB: Bank of Jiujiang, BOCD: Bank of Chengdu, HSB: Huishang Bank, BQD: Bank Of Qingdao, BOJ: Bank of Jiangsu, TCCB: Tianjin City Commercial Bank, WJRCB: Jiangsu Wujiang Rural Commercial Bank, CRCB: Chongqing Rural Commercial Bank, CSRCB: Jiangsu Changshu Rural Commercial Bank, ZRCB: Jiangsu Zhangjiagang Rural Commercial Bank, GRCB: Guangzhou Rural Commercial Bank).

4.3.2. Cost Efficiency (CE), Technology Gap Ratio (TGR), and Meta-Frontier Cost Efficiency (MCE)

CE, TGR, and MCE are the most widely used measures of efficiency of firms. The descriptive statistics of the CE, TGR and MCE for Chinese banks are displayed in Table 6. CE measures how well a bank performs relative to the "best-practice" bank under the same conditions, or how close it is to the minimum cost level [25]. We observed that the Chinese banks are operating at a very high level of cost efficiency with an average CE value of 0.982. This result agreed with the findings of Hsiao et al. [22], which indicated that the cost efficiency of Chinese banks had been converging and getting closer to one by the end of 2012. As our sample period was from 2013 to 2018, it is no wonder that the high CEs of Chinese banks were noticed. It is worth noting that the estimated CEs were not comparable among different bank groups. However, the efficiencies among banks of different ownerships could be compared based on MCE [54], which will be explained in detail later.

TGR measures the distance between group frontier technology and meta-frontier technology. A larger TGR value indicates that more advanced technologies were utilized by the group, such that the group's cost frontier is closer to the meta-cost frontier [54]. The average TGR for the four groups of banks ranged from 0.857 (for RCBs) to 0.946 (for SOCBs), with a total mean of 0.899 for the whole sample. The six state-owned banks took on the most superior technology, followed by CCBs and JSCBs. On the contrary, the six RCBs, as a group, were found to acquire least sophisticated technology, indicated by their group cost frontier deviating farthest from the meta-frontier. Banks in rural areas have short development times, with obvious geographical restrictions, and encounter challenges of backward information technology, such as a lack of high-speed wireless networks, big data, and cloud computing facilities. The strategies of credit risk management and services of online banking have also

been shown to be lacking in RCBs [55]. Thus, an obvious technology gap exists between RCBs and other banks.

Table 6. Summary statistics of Cost Efficiency, Technology Gap Ratio, and Meta-frontier Cost Efficiency.

Groups	Entire Sample	SOCBs	JSCBs	CCBs	RCBs
Cost Efficiency (CE)					
Mean	0.982	0.976	0.984	0.987	0.974
S.D.	0.013	0.010	0.011	0.012	0.017
Median	0.984	0.977	0.985	0.990	0.977
Min	0.945	0.961	0.960	0.947	0.945
Max	1.000	0.997	0.999	1.000	0.999
Technology Gap Ratio (TGR)					
Mean	0.899	0.946	0.887	0.906	0.857
S.D.	0.069	0.028	0.073	0.065	0.072
Median	0.909	0.949	0.874	0.916	0.857
Min	0.711	0.893	0.735	0.711	0.733
Max	0.999	0.995	0.999	0.998	0.998
Meta-frontier Cost Efficiency (MCE)					
Mean	0.883	0.923	0.872	0.893	0.834
S.D.	0.063	0.023	0.065	0.059	0.061
Median	0.897	0.930	0.860	0.908	0.831
Min	0.702	0.868	0.733	0.702	0.727
Max	0.981	0.964	0.980	0.981	0.979
Obs.	222	36	60	90	36

MCE can help to figure out efficiency of individual banks and, consequently, to compare banks. The average MCE of the entire sample was 0.883. SOCBs were the most efficient banks among four groups, with an average MCE of 0.923, followed by CCBs (0.893) and JSCBs (0.872); while RCBs were the least efficient banks, with an average MCE of 0.834. These results were in line with many previous studies focused on analyzing cost efficiency of Chinese banks, such as Chen et al. [28]. Nevertheless, our findings were contrary to Lee and Huang [56], who showed that JSCBs and RCBs were more cost efficient compared with CCBs and the four biggest SOCBs. The different conclusions may have been due to the different sample of banks selected, different methodologies applied and different periods covered.

4.3.3. Brief Summaries

Our findings on various bank behaviors can be summarized as follows: First, the findings by LI indicated that the competition in the Chinese banking industry mainly existed in state-owned banks and joint stock banks. Moreover, the market power of Chinese banks had increased in recent years and gaps in market power amongst banks were declining gradually. Second, the presence of economies of scale was found in CCBs and RCBs, while diseconomies of scale occurred in SOCBs and JSCBs. Third, Chinese banks in our sample, in general, were operating with a high level of cost efficiency during 2013 to 2018. The findings from TGR and MCE were consistent too: SOCBs were found to have the most superior technology leading to the highest MCE followed by CCBs and JSCBs; while TGR and MCE of RCBs were lowest among the four groups of banks.

5. Conclusions

In this study, we proposed a simultaneous SFM with dependent error components by using copula functions. The flexibility of this model is that it allows for dependence between the random noise and inefficiency components of individual SFM as well as dependence between the composite

errors of simultaneous stochastic frontier equations. We first verified reasonability of allowing for such dependence between statistical noise and inefficiency components when estimating single-equation SFM and/or simultaneous SFM using two simulation studies. The results confirmed that the copula-based SFM outperforms conventional SFM as the two error components, i.e., random noise and inefficiency components, are correlated in practice; while, for the case of simultaneous equations, it was demonstrated that ignoring dependence between random noise and inefficiency components could result in biased estimates. We then applied our developed model to measure performance of the Chinese banking industry. Empirical analysis again confirmed that our simultaneous SFM with dependent error components was superior to the other two models. Finally, we estimated market power, economies of scale, cost efficiency and technology gap ratios of 37 Chinese banks covering the period 2013–2018 based on the estimation results from the simultaneous SFM with dependent error components.

Results from the empirical application reveal that Chinese banks operated at a high level of cost efficiency and there is high level of competition amongst state-owned banks and joint stock banks, implying low market power. On the other hand, high market power was enjoyed by CCBs and RCBs. Different groups of banks demonstrated economies of scales as well as diseconomies of scales. The state-owned banks acquired the highest level of technologies which enabled them to operate at the highest level of cost efficiency.

The following policy implications can be drawn from the empirical results of the study. First, Chinese banks experiencing diseconomies of scale should consider reducing their operation size to remain competitive in the global financial market. Second, there is a need to acquire superior and advanced technologies in operation, which will allow reducing efficiency gaps of Chinese banks. A suit of advanced technologies is available within the banking sector, which the lagging banks should explore and adopt in order to remain at the top of their game in a globally competitive financial market.

The main contribution of our study is that it provides a new method to estimate simultaneous SFM with least restrictive assumptions, which is a valuable addition to the existing literature on SFM developments. Our study could serve as a useful reference for future studies, both in terms of applying innovative research methods and/or empirical applications with least restrictive assumptions. In terms of directions for future research, first, future research could focus on incorporating other copula families to specify dependence between random error and inefficiency components instead of Gaussian or Frank copulas used in our study. Second, future research can also consider other marginal distributions for the inefficiency terms instead of half-normal or exponential distributions applied in our study. Third, solving problems of possible serial autocorrelation and heteroskedasticity in the application of our proposed model for panel data could also be a useful future research direction.

Author Contributions: Conceptualization, J.L., M.W., J.M. and S.S.; Data curation, M.W. and J.M.; Methodology, J.L. and M.W.; Software, J.L. and M.W.; Supervision, J.L. and S.S.; Writing—original draft, M.W.; Writing—review & editing, J.L., S.R. and M.W. All authors have read and agreed to the published version of the manuscript.

References

1. Fang, J.; Lau, C.-K.M.; Lu, Z.; Tan, Y.; Zhang, H. Bank performance in China: A perspective from bank efficiency, risk-taking and market competition. *Pac. Basin Financ. J.* **2019**, *56*, 290–309. [CrossRef]
2. Bos, J.W.; Kolari, J.W.; Van Lamoen, R.C. Competition and innovation: Evidence from financial services. *J. Bank. Financ.* **2013**, *37*, 1590–1601. [CrossRef]

3. Riaz, M. In competency aspects of microfinance industry: Via SFA approach. *J. Econ. Behav. Stud.* **2015**, *7*, 1–12. [CrossRef]

4. Bos, J.W.; Millone, M. Practice what you preach: Microfinance business models and operational efficiency. *World Dev.* **2015**, *70*, 28–42. [CrossRef]

5. Benedetti, I.; Branca, G.; Zucaro, R. Evaluating input use efficiency in agriculture through a stochastic frontier production: An application on a case study in Apulia (Italy). *J. Clean. Prod.* **2019**, *236*, 117609. [CrossRef]

6. Anik, A.R.; Salam, M.A. Assessing and explaining vegetable growers' efficiency in the south-eastern hilly districts of Bangladesh. *J. Asia Pac. Econ.* **2017**, *22*, 680–695. [CrossRef]

7. Wang, P.; Deng, X.; Zhang, N.; Zhang, X. Energy efficiency and technology gap of enterprises in Guangdong province: A meta-frontier directional distance function analysis. *J. Clean. Prod.* **2019**, *212*, 1446–1453. [CrossRef]

8. Sun, J.; Du, T.; Sun, W.; Na, H.; He, J.; Qiu, Z.; Yuan, Y.; Li, Y. An evaluation of greenhouse gas emission efficiency in China's industry based on SFA. *Sci. Total Environ.* **2019**, *690*, 1190–1202. [CrossRef]

9. Lai, H.-P.; Huang, C.J. Maximum likelihood estimation of seemingly unrelated stochastic frontier regressions. *J. Product. Anal.* **2013**, *40*, 1–14. [CrossRef]

10. Fernández, C.; Koop, G.; Steel, M. A Bayesian analysis of multiple-output production frontiers. *J. Econom.* **2000**, *98*, 47–79. [CrossRef]

11. Ferreira, J.T.; Steel, M.F. Model comparison of coordinate-free multivariate skewed distributions with an application to stochastic frontiers. *J. Econom.* **2007**, *137*, 641–673. [CrossRef]

12. Carta, A.; Steel, M.F. Modelling multi-output stochastic frontiers using copulas. *Comput. Stat. Data Anal.* **2012**, *56*, 3757–3773. [CrossRef]

13. Huang, T.-H.; Chiang, D.-L.; Chao, S.-W. A new approach to jointly estimating the Lerner index and cost efficiency for multi-output banks under a stochastic meta-frontier framework. *Q. Rev. Econ. Financ.* **2017**, *65*, 212–226. [CrossRef]

14. Huang, T.-H.; Hu, C.-N.; Chang, B.-G. Competition, efficiency, and innovation in Taiwan's banking industry—An application of copula methods. *Q. Rev. Econ. Financ.* **2018**, *67*, 362–375. [CrossRef]

15. Bandyopadhyay, D.; Das, A. On measures of technical inefficiency and production uncertainty in stochastic frontier production model with correlated error components. *J. Product. Anal.* **2006**, *26*, 165–180. [CrossRef]

16. Smith, M.D. Stochastic frontier models with dependent error components. *Econom. J.* **2008**, *11*, 172–192. [CrossRef]

17. El Mehdi, R.; Hafner, C.M. Inference in stochastic frontier analysis with dependent error terms. *Math. Comput. Simul.* **2014**, *102*, 104–116. [CrossRef]

18. Wiboonpongse, A.; Liu, J.; Sriboonchitta, S.; Denoeux, T. Modeling dependence between error components of the stochastic frontier model using copula: Application to intercrop coffee production in Northern Thailand. *Int. J. Approx. Reason.* **2015**, *65*, 34–44. [CrossRef]

19. Sriboonchitta, S.; Liu, J.; Wiboonpongse, A.; Denoeux, T. A double-copula stochastic frontier model with dependent error components and correction for sample selection. *Int. J. Approx. Reason.* **2017**, *80*, 174–184. [CrossRef]

20. Lin, H.L.; Tsao, C.C.; Yang, C.H. Bank reforms, competition and efficiency in China's banking system: Are small city bank entrants more efficient? *China World Econ.* **2009**, *17*, 69–87. [CrossRef]

21. Fungáčová, Z.; Pessarossi, P.; Weill, L. Is bank competition detrimental to efficiency? Evidence from China. *China Econ. Rev.* **2013**, *27*, 121–134. [CrossRef]

22. Hsiao, C.; Yan, S.; Wenlong, B. Evaluating the effectiveness of China's financial reform—The efficiency of China's domestic banks. *China Econ. Rev.* **2015**, *35*, 70–82. [CrossRef]

23. Hou, X.; Wang, Q.; Zhang, Q. Market structure, risk taking, and the efficiency of Chinese commercial banks. *Emerg. Mark. Rev.* **2014**, *20*, 75–88. [CrossRef]

24. Zhu, N.; Wu, Y.; Wang, B.; Yu, Z. Risk preference and efficiency in Chinese banking. *China Econ. Rev.* **2019**, *53*, 324–341. [CrossRef]

25. Berger, A.N.; Hasan, I.; Zhou, M. Bank ownership and efficiency in China: What will happen in the world's largest nation? *J. Bank. Financ.* **2009**, *33*, 113–130. [CrossRef]

26. Chen, Z.; Matousek, R.; Wanke, P. Chinese bank efficiency during the global financial crisis: A combined approach using satisficing DEA and Support Vector Machines. *North Am. J. Econ. Financ.* **2018**, *43*, 71–86. [CrossRef]

27. Fungáčová, Z.; Klein, P.-O.; Weill, L. Persistent and transient inefficiency: Explaining the low efficiency of Chinese big banks. *China Econ. Rev.* **2020**, *59*, 101368. [CrossRef]

28. Chen, X.; Skully, M.; Brown, K. Banking efficiency in China: Application of DEA to pre- and post-deregulation eras: 1993–2000. *China Econ. Rev.* **2005**, *16*, 229–245. [CrossRef]

29. Wang, K.; Huang, W.; Wu, J.; Liu, Y.-N. Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. *Omega* **2014**, *44*, 5–20. [CrossRef]

30. Jiang, H.; He, Y. Applying data envelopment analysis in measuring the efficiency of Chinese listed banks in the context of macroprudential framework. *Mathematics* **2018**, *6*, 184. [CrossRef]

31. Yin, H.; Yang, J.; Mehran, J. An empirical study of bank efficiency in China after WTO accession. *Glob. Financ. J.* **2013**, *24*, 153–170. [CrossRef]

32. Silva, T.C.; Tabak, B.M.; Cajueiro, D.O.; Dias, M.V.B. A comparison of DEA and SFA using micro-and macro-level perspectives: Efficiency of Chinese local banks. *Phys. A: Stat. Mech. Appl.* **2017**, *469*, 216–223. [CrossRef]

33. Huang, T.-H.; Lin, C.-I.; Chen, K.-C. Evaluating efficiencies of Chinese commercial banks in the context of stochastic multistage technologies. *Pac. Basin Financ. J.* **2017**, *41*, 93–110. [CrossRef]

34. Liu, J.; Wang, M.; Sriboonchitta, S. Examining the Interdependence between the Exchange Rates of China and ASEAN Countries: A Canonical Vine Copula Approach. *Sustainability* **2019**, *11*, 5487. [CrossRef]

35. Shan, Q.; Wongyang, T.; Wang, T.; Tasena, S. A measure of mutual complete dependence in discrete variables through subcopula. *Int. J. Approx. Reason.* **2015**, *65*, 11–23. [CrossRef]

36. Kreinovich, V.; Nguyen, H.T.; Sriboonchitta, S.; Kosheleva, O. Why copulas have been successful in many practical applications: A theoretical explanation based on computational efficiency. In *International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making*; Springer: Cham, Switzerland, 2015; pp. 112–125.

37. Trivedi, P.; Zimmer, D. A note on identification of bivariate copulas for discrete count data. *Econometrics* **2017**, *5*, 10. [CrossRef]

38. Wei, Z.; Kim, D. On multivariate asymmetric dependence using multivariate skew-normal copula-based regression. *Int. J. Approx. Reason.* **2018**, *92*, 376–391. [CrossRef]

39. Zhu, X.; Wang, T.; Zhang, X.; Wang, L. Comparisons on Measures of Asymmetric Associations. In *Proceedings of International Econometric Conference of Vietnam*; Springer: Cham, Switzerland, 2019; pp. 185–197.

40. Wu, C.-C.; Chung, H.; Chang, Y.-H. The economic value of co-movement between oil price and exchange rate using copula-based GARCH models. *Energy Econ.* **2012**, *34*, 270–282. [CrossRef]

41. Sriboonchitta, S.; Nguyen, H.T.; Wiboonpongse, A.; Liu, J. Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. *Int. J. Approx. Reason.* **2013**, *54*, 793–808. [CrossRef]

42. Wei, Z.; Kim, S.; Choi, B.; Kim, D. Multivariate Skew Normal Copula for Asymmetric Dependence: Estimation and Application. *Int. J. Inf. Technol. Decis. Mak.* **2019**, *18*, 365–387. [CrossRef]

43. Fall, F.; Akim, A.-M.; Wassongma, H. DEA and SFA research on the efficiency of microfinance institutions: A meta-analysis. *World Dev.* **2018**, *107*, 176–188. [CrossRef]

44. Coelli, T.J. *A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*; CEPA Working Papers; University of New England: Armidale, Australia, 1996.

45. Morris, T.P.; White, I.R.; Crowther, M.J. Using simulation studies to evaluate statistical methods. *Stat. Med.* **2019**, *38*, 2074–2102. [CrossRef] [PubMed]

46. Greene, W. A stochastic frontier model with correction for sample selection. *J. Product. Anal.* **2010**, *34*, 15–24. [CrossRef]

47. Shamshur, A.; Weill, L. Does bank efficiency influence the cost of credit? *J. Bank. Financ.* **2019**, *105*, 62–73. [CrossRef]

48. Spierdijka, L.; Zaourasa, M. Measuring banks' market power in the presence of economies of scale: A scale-corrected Lerner index. *J. Bank. Financ.* **2018**, *87*, 40–48. [CrossRef]

49. Tan, Y. The impacts of risk and competition on bank profitability in China. *J. Int. Financ. Mark. Inst. Money* **2016**, *40*, 85–110. [CrossRef]

50. Druska, V.; Horrace, W.C. Generalized moments estimation for spatial panel data: Indonesian rice farming. *Am. J. Agric. Econ.* **2004**, *86*, 185–198. [CrossRef]

51. Berger, A.N.; Humphrey, D.B. *Bank Scale Economies, Mergers, Concentration, and Efficiency: The US Experience*; Working Papers 94-25; Wharton School Center for Financial Institutions, University of Pennsylvania: Philadelphia, PA, USA, 1994.

52. Barros, C.P.; Ferreira, C.; Williams, J. Analysing the determinants of performance of best and worst European banks: A mixed logit approach. *J. Bank. Financ.* **2007**, *31*, 2189–2203. [CrossRef]

53. Athanasoglou, P.P.; Brissimis, S.N.; Delis, M. Bank-specific, industry-specific and macroeconomic determinants of bank profitability. *J. Int. Financ. Mark. Inst. Money* **2008**, *18*, 121–136. [CrossRef]

54. Lee, C.-C.; Huang, T.-H. Cost efficiency and technological gap in Western European banks: A stochastic metafrontier analysis. *Int. Rev. Econ. Financ.* **2017**, *48*, 161–178. [CrossRef]

55. Wang, Y.; Wang, W.; Wang, J. Credit risk management framework for rural commercial banks in China. *J. Financ. Risk Manag.* **2017**, *6*, 48. [CrossRef]

56. Lee, C.C.; Huang, T.H. What causes the efficiency and the technology gap under different ownership structures in the Chinese banking industry? *Contemp. Econ. Policy* **2019**, *37*, 332–348. [CrossRef]