Journal of Statistical Theory and Practice (2022) 16:20
https://doi.org/10.1007/s42519-022-00249-9

ORIGINAL ARTICLE

Multivariate Skew Normal-Based Stochastic Frontier Models

Xiaonan Zhu[1] · Zheng Wei[2] · Tonghui Wang[3]

Accepted: 25 February 2022

Abstract

In this paper, a new multivariate skew normal-based stochastic frontier model (SFM) is developed to model the relationship between technical inefficiency and random noises. The proposed model is a flexible extension of the classical normal–half-normal SFM and does not require the data are independent and identically distributed. The new proposed model accommodates both positive and negative skewness and provides an alternative solution of 'wrong skewness' problems from the classical SFM. Maximum likelihood estimators of parameters are provided. The simulation study and a real data example are given for illustration of our results.

Keywords Multivariate skew normal distribution · Stochastic frontier model · Technical efficiency · Wrong skewness problem

1 Introduction

The stochastic frontier model (SFM) is the most useful approach to study productivity and efficiency of a cross section of firms independently proposed by Aigner et al. [1] and Meeusen and Broeck [2]. The basic model can be written as:

✉ Zheng Wei
zheng.wei@maine.edu

Xiaonan Zhu
xzhu7@una.edu

Tonghui Wang
twang@nmsu.edu

1 Department of Mathematics, University of North Alabama, Stevens Hall 309D, Florence 35632, AL, USA

2 Department of Mathematics and Statistics, University of Maine, 5752 Neville Hall Room 330, Orono, ME 04469-5752, USA

3 Department of Mathematical Sciences, New Mexico State University, Science Hall 254, Las Cruces 88003-8001, NM, USA

Published online: 22 March 2022

$$Y_i = \boldsymbol{X}_i^\top \boldsymbol{\beta} + V_i - U_i, \quad i = 1, \ldots, n. \tag{1}$$

The regression part of SFM is the production frontier of the ith firm. The dependent variable Y_i denotes the output and the vector $\boldsymbol{X}_i$ is inputs. The error term V_i represents stochastic effects outside the control of the ith firm, and the ith firm possesses an inefficiency term $U_i \geq 0$ [3], In the standard SFM, people usually assume that $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ and $\boldsymbol{U} = (U_1, \ldots, U_n)^\top$ are independent such that V_i's are independent and identically distributed (i.i.d.) with $V_i \sim \mathcal{N}(0, \sigma_v^2)$ and U_i's are i.i.d. on $\mathbb{R}^+$.

It is well known [4] that the third moment of $\mathcal{E}_i = V_i - U_i$ is given by

$$E\left[(\mathcal{E}_i - E[\mathcal{E}_i])^3\right] = -E\left[(U_i - E(U_i))^3\right],$$

so that a positive skewness of U_i implies a negative skewness for $\mathcal{E}_i$. This restriction in SFM could be incompatible with the real data when residuals shows positive skewness. This is known as the wrong skewness issue in SFM literature [5, 6]. The wrong skewness problem implies that the estimated variance of U is zero, which means full efficiency for all firms across the whole industry.

In order to solve the wrong skewness problem, Smith [7] relaxed the independence assumption between $\boldsymbol{V}$ and $\boldsymbol{U}$ by applying the symmetric copula models in SFM. Bonanno [8] applied the FGM copula for the dependence between $\boldsymbol{V}$ and $\boldsymbol{U}$. More recently, the copula-based SFM was extended by the asymmetric copula, the skew normal copula [9].

In this paper, we propose a multivariate skew normal-based SFM (MSN-SFM). It assumes that components of the error term $\boldsymbol{V}$ in SFM are identically skew normally distributed such that their joint distribution is a multivariate skew normal distribution defined by Azzalini and Dalla Valle [10]. The components of the inefficiency term $\boldsymbol{U}$ are i.i.d. half normally distributed. For simplicity, we assume that $\boldsymbol{U}$ and $\boldsymbol{V}$ are independent. The main advantages of the new proposed model are as follows. (i) In reality, there is no way to justify that Y_i's are independent. Under the assumptions of our model, $\mathcal{E}_i$'s are not independent but identically distributed when skewness parameter of V_i is not zero. (ii) The skewness in the composite error $\mathcal{E}_i$ allows both positive and negative skewness; thus, it provides an alternative solution when the wrong skewness appears. (iii) Recently, it has been shown that the mean square error of the location estimator under the dependence assumption is smaller than that under the independent assumption [11].

Simulation studies on the performance of the proposed MSN-SFM are given with comparisons of the classical normal–half-normal SFM and i.i.d. skew normal–half-normal SFM (SN-SFM). A real data application with the wrong skewness under the classical SFM shows that the proposed MSN-SFM performs better than other two models. Our proposed MSN-SFM provides nonzero inefficiency estimates among firms.

This paper is organized as follows. The brief review of skew normal families as preliminaries are given in Sect. 2. The main results of the MSN-SFM and SN-SFM are provided in Sect. 3, and the skewness of the composite in new proposed MSN-SFM is derived. The firm-specific efficiency is discussed in Sect. 4. The maximum likelihood estimation of parameters and the simulation study are included in Sect. 5. In Sect. 6, real data with wrong skewness are analyzed under new proposed models and SFM for the illustration of our main results. Conclusion and further works are given in Sect. 7.

2 Preliminaries

In this paper, $M_{n\times m}$ denotes the set of all $n \times m$ matrices over the real field $\mathbb{R}$ and $\mathbb{R}^n = M_{n\times 1}$. For any $\boldsymbol{B} \in M_{n\times m}$, $\boldsymbol{B}^\top$ is the transpose of $\boldsymbol{B}$. $\boldsymbol{I}_n \in M_{n\times n}$ is the identity matrix. $\boldsymbol{a}_n$ is the column vector $(a, \cdots, a)^\top \in \mathbb{R}^n$. $\boldsymbol{J}_n = \mathbf{1}_n\mathbf{1}_n^\top$. Let $\boldsymbol{b} = (b_1, \ldots, b_n)^\top \in \mathbb{R}^n$, $\mathrm{diag}(b_1, \ldots, b_n)$ is the diagonal matrix of $\boldsymbol{b}$. $\phi_n(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\Phi_n(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ are the probability density function (PDF) and cumulative distribution function (CDF), respectively, of an n-dimensional normal distribution with the mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\boldsymbol{\Sigma}$, simply, $\phi_n(\cdot\,; \boldsymbol{\Sigma})$ and $\Phi_n(\cdot\,; \boldsymbol{\Sigma})$ for the case when $\boldsymbol{\mu} = \mathbf{0}_n$. Also, $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the PDF and CDF of the univariate standard normal distribution.

The multivariate skew normal (MSN) distribution was introduced by Azzalini and Dalla Valle [10]. Its extensions and applications have been studied by many researchers [12–14]. Consider the $n+1$-dimensional multivariate normal distribution given as follows:

$$\begin{pmatrix} \boldsymbol{Z}_0 \\ Z_1 \end{pmatrix} \sim \mathcal{N}_{n+1}\left(\mathbf{0}, \begin{pmatrix} \Psi, & \mathbf{0} \\ \mathbf{0}, & 1 \end{pmatrix}\right),$$

$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top$ with $-1 \le \delta_1, \ldots, \delta_n \le 1$, and $\boldsymbol{D}_\delta = \mathrm{diag}\left(\sqrt{1-\delta_1^2}, \ldots, \sqrt{1-\delta_n^2}\right)$. Defining

$$\boldsymbol{Z} = \boldsymbol{D}_\delta \boldsymbol{Z}_0 + \boldsymbol{\delta}|Z_1|,$$

then $\boldsymbol{Z}$ has an n-dimensional skew normal distribution with parameters $(\bar{\Omega}, \boldsymbol{\alpha})$ related to $\boldsymbol{\delta}$ and Ψ as follows:

$$\boldsymbol{\lambda} = \boldsymbol{D}_\delta^{-1}\boldsymbol{\delta}, \quad \bar{\Omega} = \boldsymbol{D}_\delta(\Psi + \boldsymbol{\lambda}\boldsymbol{\lambda}^\top)\boldsymbol{D}_\delta, \quad \boldsymbol{\alpha} = \left(1 + \boldsymbol{\lambda}^\top\Psi^{-1}\boldsymbol{\lambda}\right)^{-1/2}\boldsymbol{D}_\delta^{-1}\Psi^{-1}\boldsymbol{\lambda},$$

denoted by $\boldsymbol{Z} \sim \mathcal{SN}_n(\bar{\Omega}, \boldsymbol{\alpha})$. Let $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\omega}\boldsymbol{Z}$, where $\boldsymbol{\mu} \in \mathbb{R}^n, \boldsymbol{\omega} = \mathrm{diag}(\omega_1, \ldots, \omega_n), \omega_i > 0, i = 1, \ldots, n$. The PDF of $\boldsymbol{Y}$ is given by

$$f(\boldsymbol{y}) = 2\phi_n(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi(\boldsymbol{\alpha}^\top\boldsymbol{\omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})), \tag{2}$$

where $\boldsymbol{\Sigma} = \boldsymbol{\omega}\bar{\Omega}\boldsymbol{\omega}$. $\boldsymbol{Y}$ is called a multivariate skew normal random vector with location parameter $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and skewness $\boldsymbol{\alpha}$, denoted by $\mathcal{SN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$. From the construction of MSN distributions, we can see that components of an MSN random vector are dependent when the skewness parameter of each component is not zero.

In order to prove our main results, we need the closed skew normal (CSN) distribution defined by González-Farías et al. [15]. A random vector $\boldsymbol{X} \in \mathbb{R}^p$ is said to have a CSN distribution with parameters $q \geq 1$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma} \in M_{p\times p}$, $\boldsymbol{D} \in M_{q\times p}$, $\boldsymbol{\nu} \in \mathbb{R}^q$, $\boldsymbol{\Delta} \in M_{q\times q}$, denoted by $\boldsymbol{X} \sim \mathcal{CSN}_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{D}, \boldsymbol{\nu}, \boldsymbol{\Delta})$, if its PDF is given by

$$f_{p,q}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{D}, \boldsymbol{\nu}, \boldsymbol{\Delta}) = C\phi_p(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi_q(\boldsymbol{D}(\boldsymbol{x} - \boldsymbol{\mu}); \boldsymbol{\nu}, \boldsymbol{\Delta}), \quad \boldsymbol{x} \in \mathbb{R}^p,$$

where $C^{-1} = \Phi_q(\boldsymbol{0}; \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^T)$.

Remark 2.1 The closed-skew normality was applied to the stochastic frontiers in [16], where the vector of random errors $\boldsymbol{V}$ is assumed to be the multivariate normal distributed. In this paper, we assume that $\boldsymbol{V}$ follows MSN with skewness vector $\boldsymbol{\lambda}$, which generalizes the results in [16] for the case where $T = 1$.

The moment generating function (MGF) of the CSN distribution, which will be utilized for constructing the MSN-based SFM, is given in the following lemma.

Lemma 2.1 [15] *Let* $\boldsymbol{X} \sim CSN_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, D, \boldsymbol{\nu}, \boldsymbol{\Delta})$*, then the MGF of* $\boldsymbol{X}$ *is,*

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \frac{\Phi_q(\boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{t}; \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^T)}{\Phi_q(\boldsymbol{0}; \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^T)} \exp\left(\boldsymbol{t}^\top\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Sigma}\boldsymbol{t}\right), \quad \boldsymbol{t} \in \mathbb{R}^p.$$

3 Multivariate Skew Normal-Based SFM and the Skewness of the Composite Error

In this section, we consider case where $\Psi = \boldsymbol{I}_n$ in Eq. (2) so that

$$\boldsymbol{V} = (V_1, \dots, V_n)^\top \sim \mathcal{SN}_n(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}),$$

where

$$\boldsymbol{\Sigma} = \sigma_v^2\boldsymbol{D}_\delta(\boldsymbol{I}_n + \boldsymbol{\lambda}\boldsymbol{\lambda}^\top) \quad \text{and} \quad \boldsymbol{\alpha} = \left(1 + \boldsymbol{\lambda}^\top\boldsymbol{\lambda}\right)^{-1/2}\boldsymbol{D}_\delta^{-1}\boldsymbol{\lambda}.$$

Thus, it is easy to show that $V_i \sim \mathcal{SN}_1(0, \sigma_v^2, \lambda_i)$, $i = 1, \dots, n$. Specially, if λ_i's are equal, then V_i's are identically distributed.

The following result is our main theorem on the joint distribution of the composite error term in SFM in Eq. (1).

Theorem 3.1 *Assume that* $\boldsymbol{V} \sim \mathcal{SN}_n(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ *given above, and* $\boldsymbol{U}$ *is distributed as the* n*-dimensional half-normal distribution with the covariance matrix* $\sigma_u^2\boldsymbol{I}_n$*, denoted as* $\mathcal{HN}_n(\boldsymbol{0}_n, \sigma_u^2\boldsymbol{I}_n)$*. Then, the joint density of* $\boldsymbol{\mathcal{E}} = \boldsymbol{V} - \boldsymbol{U}$ *in Eq. (1) is*

$$f_{\mathcal{E}}(\boldsymbol{\varepsilon}) = C\phi_n\left(\boldsymbol{\varepsilon}; \boldsymbol{0}_n, \boldsymbol{\Sigma} + \sigma_u^2\boldsymbol{I}_n\right)\Phi_{n+1}\left(\boldsymbol{D}\boldsymbol{\varepsilon}; \boldsymbol{0}_{n+1}, \boldsymbol{\Delta}\right),$$

where

$$
\begin{aligned}
C^{-1} &= \Phi_{n+1}\left(\mathbf{0}_{n+1}; \mathbf{0}_{n+1}, \boldsymbol{\Delta} + \boldsymbol{D}(\boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)\boldsymbol{D}^\top\right), \\
\boldsymbol{D} &= \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1}, \\
\boldsymbol{\Delta} &= \boldsymbol{I}_{n+1} - \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1}(\boldsymbol{b}, -\sigma_u \boldsymbol{I}_n), \\
\boldsymbol{b} &= \left(\sigma_v^2 + \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\alpha}\right)^{-1/2} \boldsymbol{\Sigma}\boldsymbol{\alpha}.
\end{aligned}
$$

Proof Since $\boldsymbol{V} \sim \mathcal{SN}_n(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, by the PDF of $\mathcal{SN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ given in Eq (2), $\boldsymbol{V} \sim \mathcal{CSN}_{n,1}\left(\mathbf{0}_n, \boldsymbol{\Sigma}, \boldsymbol{\alpha}^\top/\sigma_v, 0, 1\right)$. Furthermore, the MGF of $\boldsymbol{U}$ is

$$
M_U(\boldsymbol{t}) = \frac{\Phi_n(\sigma_u \boldsymbol{t}; \mathbf{0}_n, \boldsymbol{I}_n)}{\Phi_n(\mathbf{0}_n; \mathbf{0}_n, \boldsymbol{I}_n)} \exp\left(\frac{1}{2}\sigma_u^2 \boldsymbol{t}^\top \boldsymbol{t}\right).
$$

Thus, the MGF of $\boldsymbol{\mathcal{E}} = \boldsymbol{V} - \boldsymbol{U}$ is given by

$$
M_{\mathcal{E}}(\boldsymbol{t}) = E\left[e^{\boldsymbol{t}^\top \boldsymbol{V}} e^{-\boldsymbol{t}^\top \boldsymbol{U}}\right] = \frac{\Phi(\boldsymbol{b}^\top \boldsymbol{t})\Phi_n(-\sigma_u \boldsymbol{t})}{\Phi(0)\Phi_n(\mathbf{0}_n)} \exp\left\{\frac{1}{2}\left(\boldsymbol{t}^\top \Omega \boldsymbol{t} + \sigma_u^2 \boldsymbol{t}^\top \boldsymbol{t}\right)\right\}.
$$

It can be shown that $\boldsymbol{\mathcal{E}}$ follows $\mathcal{CSN}_{n,n+1}(\mathbf{0}_n, \Omega + \sigma_u^2 \boldsymbol{I}_n, \boldsymbol{D}, \mathbf{0}_{n+1}, \boldsymbol{\Delta})$ with

$$
\begin{aligned}
\boldsymbol{D} &= \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1}, \quad \boldsymbol{b} = \left(\sigma_v^2 + \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\alpha}\right)^{-1/2} \boldsymbol{\Sigma}\boldsymbol{\alpha}, \\
\boldsymbol{\Delta} &= \boldsymbol{I}_{n+1} - \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1}(\boldsymbol{b}, -\sigma_u \boldsymbol{I}_n).
\end{aligned}
$$

The desired result follows. □

If we assume that components of $\boldsymbol{\mathcal{E}}$ are identically distributed, then the distribution of $\boldsymbol{\mathcal{E}}$ is given by the following results.

Corollary 3.1 *Let* $\boldsymbol{U} \sim \mathcal{HN}_n(0, \sigma_u^2 \boldsymbol{I}_n)$ *and* $\boldsymbol{V} \sim \mathcal{SN}_n(0, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, *where*

$$
\boldsymbol{\Sigma} = \sigma_v^2\left[(1-\delta^2)\boldsymbol{I}_n + \delta^2 \boldsymbol{J}_n\right] \quad \text{and} \quad \boldsymbol{\alpha} = \frac{\delta}{\sqrt{1+(n-2)\delta^2 - (n-1)\delta^4}}\mathbf{1}_n.
$$

(i) The joint density of $\boldsymbol{\mathcal{E}} = \boldsymbol{V} - \boldsymbol{U}$ *in Eq.* (1) *is*

$$
f_{\mathcal{E}}(\boldsymbol{\varepsilon}) = C\phi_n\left(\boldsymbol{\varepsilon}; \mathbf{0}_n, \boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n\right) \Phi_{n+1}\left(\boldsymbol{D}\boldsymbol{\varepsilon}; \mathbf{0}_{n+1}, \boldsymbol{\Delta}\right), \tag{3}
$$

where

$$\begin{aligned}
C^{-1} &= \Phi_{n+1}\left(\mathbf{0}_{n+1}; \mathbf{0}_{n+1}, \mathbf{\Delta} + \boldsymbol{D}(\mathbf{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)\boldsymbol{D}^\top\right) \\
\boldsymbol{D} &= \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\mathbf{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1}, \\
\mathbf{\Delta} &= \boldsymbol{I}_{n+1} - \begin{pmatrix} \boldsymbol{b}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} (\mathbf{\Sigma} + \sigma_u^2 \boldsymbol{I}_n)^{-1} (\boldsymbol{b}, -\sigma_u \boldsymbol{I}_n) \\
\boldsymbol{b} &= \left(\sigma_v^2 + \boldsymbol{\alpha}^\top \mathbf{\Sigma} \boldsymbol{\alpha}\right)^{-1/2} \mathbf{\Sigma}\boldsymbol{\alpha}.
\end{aligned}$$

(ii) The distribution of the average of $\boldsymbol{\mathcal{E}}$, $\bar{\boldsymbol{\mathcal{E}}} = \frac{1}{n}\mathbf{1}_n^\top \boldsymbol{\mathcal{E}}$, is $\bar{\boldsymbol{\mathcal{E}}} \sim \mathcal{CSN}_{1,n+1}\left(0, \sigma_{\bar{\varepsilon}}^2, \boldsymbol{D}_{\bar{\varepsilon}}, \mathbf{0}_{n+1}, \mathbf{\Delta}_{\bar{\varepsilon}}\right)$, where

$$\begin{aligned}
\sigma_{\bar{\varepsilon}}^2 &= \frac{1}{n}\left[\sigma_v\left(1 - \delta^2\right) + \sigma_u^2\right] + \sigma_v^2 \delta^2 \\
\boldsymbol{D}_{\bar{\varepsilon}} &= \frac{1}{n\sigma_{\bar{\varepsilon}}^2} \begin{pmatrix} \boldsymbol{b}_{\bar{\varepsilon}}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} \mathbf{1}_n, \\
\mathbf{\Delta}_{\bar{\varepsilon}} &= \boldsymbol{I}_{n+1} - \frac{1}{n^2\sigma_{\bar{\varepsilon}}^2} \begin{pmatrix} \boldsymbol{b}_{\bar{\varepsilon}}^\top \\ -\sigma_u \boldsymbol{I}_n \end{pmatrix} \boldsymbol{J}_n (\boldsymbol{b}_{\bar{\varepsilon}}, -\sigma_u \boldsymbol{I}_n), \\
\boldsymbol{b}_{\bar{\varepsilon}} &= \left(\sigma_v^2 + \boldsymbol{\alpha}^\top \mathbf{\Sigma} \boldsymbol{\alpha}\right)^{-1/2} \mathbf{\Sigma}\boldsymbol{\alpha}.
\end{aligned}$$

The density curves of $\bar{\boldsymbol{\mathcal{E}}}$ with various of σ_u, σ_v, $\lambda = \delta/\sqrt{1-\delta^2}$ and sample size n are given in Fig. 1. From the figure, we can see that the skewness of $\bar{\boldsymbol{\mathcal{E}}}$ increases as λ or sample size n increases.

For the case where V_i's are i.i.d. as $\mathcal{SN}_1(0, \sigma_v^2, \lambda)$, we have an alternative model called skew normal SFM (SN-SFM) given in the following theorem.

Theorem 3.2 *[6] Consider i.i.d. components of $\boldsymbol{V}$, i.e., $V_i \sim \mathcal{SN}_1(0, \sigma_v^2, \lambda)$ and $U_i \sim \mathcal{HN}_1(0, \sigma_u^2)$. The PDF of $\mathcal{E}_i = V_i - U_i$ is*

$$f_{\mathcal{E}_i}(\varepsilon_i) = C_0 \phi\left(\varepsilon_i; 0, \sigma_\varepsilon^2\right) \Phi_2\left(\boldsymbol{D}\varepsilon_i; \mathbf{0}_2, \mathbf{\Delta}\right), \quad \varepsilon_i \in \mathbb{R},$$

where $C_0^{-1} = \Phi_2\left(\mathbf{0}_2; \mathbf{0}_2, \mathbf{\Delta} + \sigma_\varepsilon^2 \boldsymbol{D}\boldsymbol{D}^\top\right)$, $\boldsymbol{D} = \frac{1}{\sigma_\varepsilon^2}(\delta\sigma_v, -\sigma_u)^\top$, and

$$\Delta = \boldsymbol{I}_2 - \frac{1}{\sigma_\varepsilon^2}\begin{pmatrix} \delta^2\sigma_v^2 & -\delta\sigma_v\sigma_u \\ -\delta\sigma_v\sigma_u & \sigma_u^2 \end{pmatrix}, \quad \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}, \quad \sigma_\varepsilon^2 = \sigma_u^2 + \sigma_v^2.$$

The proof is similar to that given in Theorem 3.1.

Next, we are going to show that our proposed models allow both positive and negative skewness.

Proposition 3.1 *Let* $\boldsymbol{U} \sim \mathcal{HN}_n(0, \sigma_u^2 \boldsymbol{I}_n)$ *and* $\boldsymbol{V} \sim \mathcal{SN}_n(0, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, *where*

$$\boldsymbol{\Sigma} = \sigma_v^2 \left[(1-\delta^2)\boldsymbol{I}_n + \delta^2 \boldsymbol{J}_n \right] \quad \text{and} \quad \boldsymbol{\alpha} = \frac{\delta}{\sqrt{1+(n-2)\delta^2 - (n-1)\delta^4}} \mathbf{1}_n.$$

The skewness of $\mathcal{E}_i = V_i - U_i$ *is given by*

$$E\left[\left(\frac{\mathcal{E}_i - E[\mathcal{E}_i]}{\sigma_\varepsilon} \right)^3 \right] = \frac{E[\mathcal{E}_i^3] - 3E[\mathcal{E}_i]E[\mathcal{E}_i^2] + 2(E[\mathcal{E}_i])^3}{(E[\mathcal{E}_i^2] - (E[\mathcal{E}_i])^2)^3}, \tag{4}$$

where the first three moments of $\mathcal{E}_i$ *are given by*

$$E(\mathcal{E}_i) = \sqrt{\frac{2}{\pi}}(\delta\sigma_v - \sigma_u), \qquad E(\mathcal{E}_i^2) = \sigma_\varepsilon^2 - \frac{4}{\pi}\delta\sigma_u\sigma_v,$$
$$E(\mathcal{E}_i^3) = \sqrt{\frac{2}{\pi}} \left(3\delta\sigma_v\sigma_\varepsilon^2 - \delta^3\sigma_v^3 - 3\sigma_u\sigma_\varepsilon^2 + \sigma_u^3 \right),$$

with $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, $\sigma_\varepsilon^2 = \sigma_v^2 + \sigma_u^2$.

Proof The proof is straightforward. □

Remark 3.1 It can be shown that the proposed MSN-SFM models accommodate both negative and positive skewness. Specifically, the skewness of $\bar{\mathcal{E}}$ given in Corollary. 3.1 as a function of δ when $\sigma_u, \sigma_v = 0.5, 1$ and $n = 1, 5, 10$, respectively, is given in Fig. 2.

Remark 3.2 From the graphs, we can see that the skewness increases as sample size n increases, so that the central limit theorem fails. This is due to the dependence assumption in the components of $\boldsymbol{V}$. This case can also be verified by Fig. 1. The distribution of $\bar{\mathcal{E}}$ given in Corollary 3.1 depends on the evaluation of Φ_{n+1}, which is difficult to calculate for large sample size n. This will be our future research topic to investigate.

4 Firm-Specific Efficiency

In this section, the inefficiency of individual firms based on the multivariate SFMs proposed in Sect. 3 is investigated. Note that V_i and U_i in SFM given in Eq. (1) are latent variables, the composite error term $\mathcal{E}$ can be obtained, but the technical efficiencies $TE = \exp(-U)$ are to be estimated. By [7, 17], the TE can be calculated by:

$$TE(\varepsilon_i; \boldsymbol{\theta}) = E[\exp(-U_i) | \mathcal{E}_i = \varepsilon_i].$$

From Theorem 3.1, the conditional distribution of U given $\mathcal{E}_i = \varepsilon_i$ and the TE, are given below.

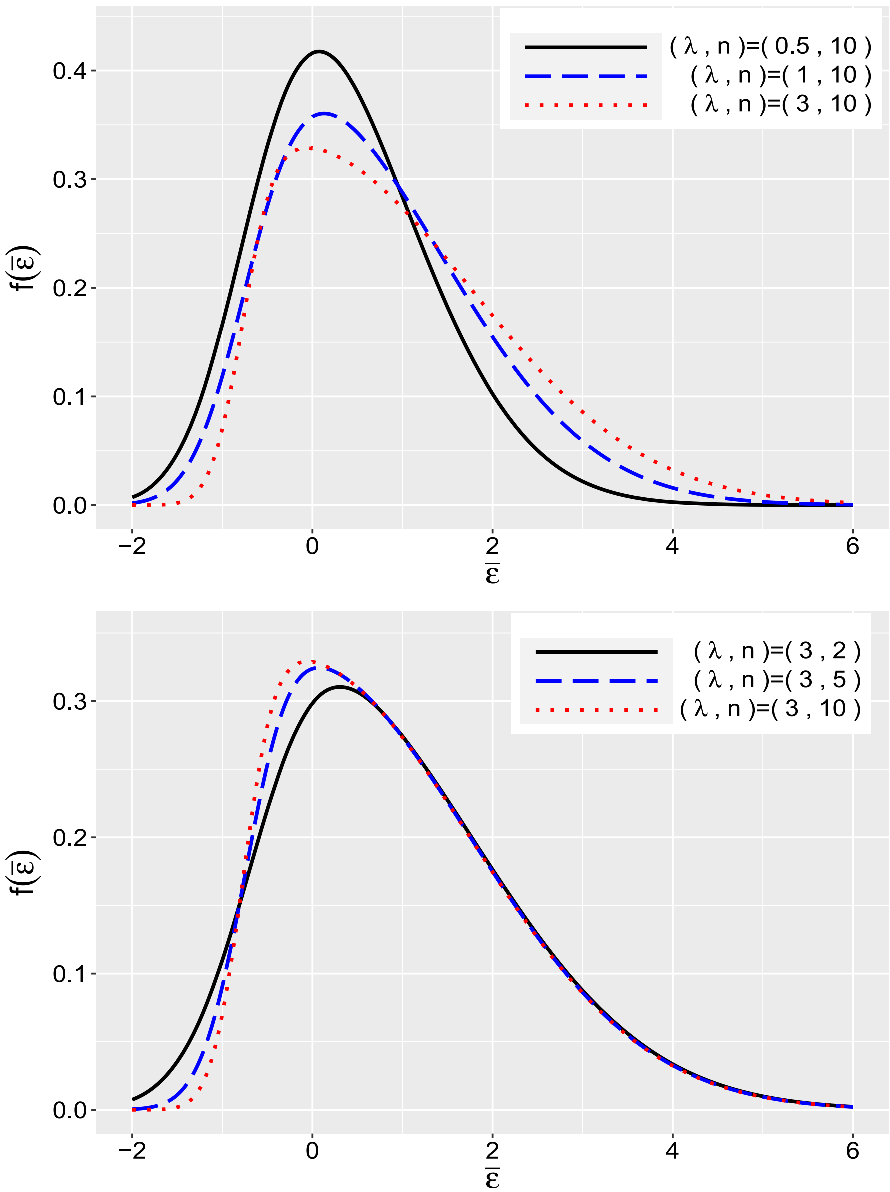

Fig. 1 Plots of $\bar{\mathcal{E}}$ for $\sigma_u = 1$ and $\sigma_v = 4$ with $\lambda = 0.5, 1, 3$ for $n = 10$ (top figure) and $n = 2, 5, 10$ for $\lambda = 3$ (bottom figure)

Proposition 4.1 *Let* $\boldsymbol{V} \sim \mathcal{MSN}_n(\mathbf{0}, \Sigma, \boldsymbol{\alpha})$ *given in Theorem 3.1, and* $\boldsymbol{U} \sim \mathcal{HN}_n(\mathbf{0}_n, \sigma_u^2 \boldsymbol{I}_n)$*, then the conditional PDF of* U_i *given* $\mathcal{E}_i = \varepsilon_i$ *is*

$$f_{U_i|\mathcal{E}_i}(u_i|\varepsilon_i) = \frac{2\phi_2\left(u_i + \varepsilon_i, u_i; (0,0)^\top, diag(\sigma_v^2, \sigma_u^2)\right)\Phi\left(\frac{\lambda}{\sigma_v}(u_i + \varepsilon_i)\right)}{C\phi(\varepsilon_i; 0, \sigma_v^2 + \sigma_u^2)\Phi_2(D\varepsilon_i; (0,0)^\top, \boldsymbol{\Delta})}, \tag{5}$$

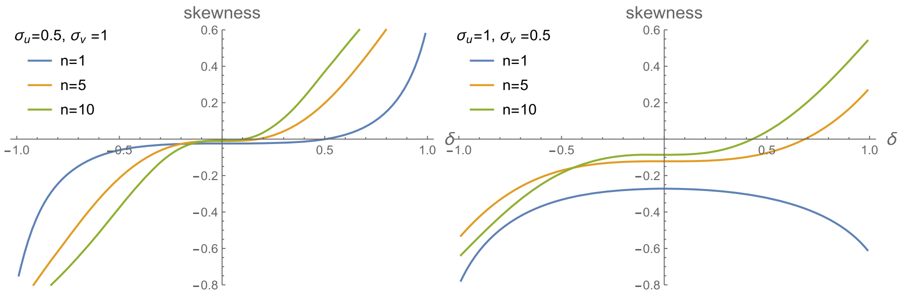

Fig. 2 The skewness of $\bar{\boldsymbol{\mathcal{E}}}$ as a function of δ when $\sigma_u, \sigma_v = 0.5, 1$, and $n = 1, 5, 10$, respectively

where $C^{-1} = \Phi_2(\mathbf{0}_2; (0, 0)^\top, \boldsymbol{\Delta} + D\boldsymbol{\Sigma}D^\top)$ *and* $\boldsymbol{\Sigma}$, D *are given in Eq.* (3)*, respectively, and*

$$TE(\varepsilon_i) = E_{U_i}[\exp(-U_i)|\mathcal{E}_i = \varepsilon_i] = \int_0^\infty \exp(-u_i) f_{U_i|\mathcal{E}_i}(u_i|\varepsilon) du_i. \tag{6}$$

Remark 4.1 The marginal PDF curves, $f_{\mathcal{E}_i}(\varepsilon_i)$, of $\mathcal{E}_i$ in Eq. (3) and their corresponding technical efficiencies, $TE(\varepsilon_i)$, in Proposition 4.1 are given in Fig. 3. The PDFs in Eq. (3) with $\sigma_v = 1$ and $\sigma_u = 1, 0.5$ are shown in the left column of Fig. 3, and the plots of the TE in Eq. (6) with $\varepsilon_i = -0.1, -0.5$ and $\sigma_v = 1$ are given in the right column. The plots in the top panel of Fig. 3 show that density and TE curves are affected by values of $\lambda = -5, -1, 0$. The plots in the bottom panel of Fig. 3 show the density and TE curves having wrong skewness (positive skewness) with $\lambda = 1$ and 5.

From Fig. 3, we can see that the values of λ play important roles in the marginal density curves of $f_{\mathcal{E}_i}(\varepsilon_i)$ and the shapes of $TE(\varepsilon_i)$: (i) $f_{\mathcal{E}_i}(\varepsilon_i)$ show both left skewed and right skewed shape depending on the values in λ and (ii) as λ increases from negative values to positive values, the shape of $f_{\mathcal{E}_i}(\varepsilon_i)$ changes from left skewed to the right skewed.

5 Maximum Likelihood Estimation and Simulation Studies

In this section, we propose the maximum likelihood estimates for parameters in both MSN-SFM and SN-SFM and derived estimates of technical efficiencies.

We assume that the frontier output is the production function of inputs in the form

$$y = f(\boldsymbol{x}; \boldsymbol{\beta}),$$

with unknown parameter vector $\boldsymbol{\beta}$. For instance the Cobb–Douglas production function has the form [7, 18]

$$y = f(\boldsymbol{x}; \boldsymbol{\beta}) = \exp(\beta_0) x_1^{\beta_1} x_2^{\beta_2} \cdots x_m^{\beta_m},$$

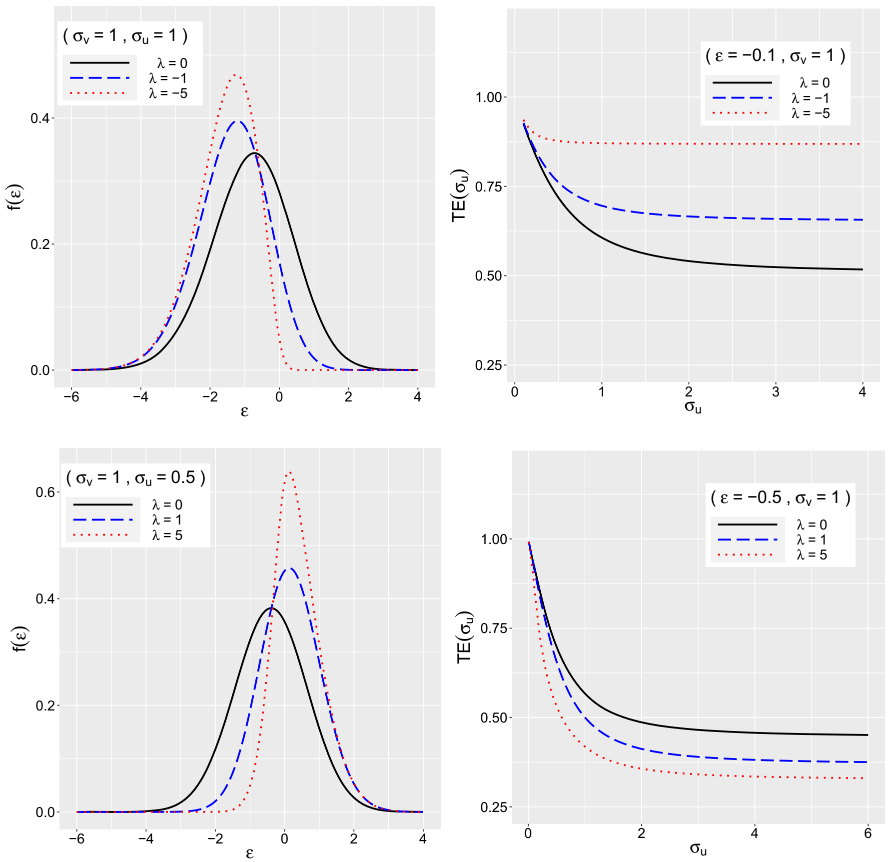

Fig. 3 Density curves of $\mathcal{E}_i$ in Eq. (3) with $\sigma_v = 1$ and $\sigma_u = 0.5, 1$ are given in left column, and curves of $TE(\varepsilon)$ in Eq. (6) with $\varepsilon_i = -0.1, -0.5$ and $\sigma_v = 1$ are given in right column. Top panel plots are the curves with $\lambda = (-5, -1, 0)$. The bottom panel plots are the curves showing wrong skewness with $\lambda = (0, 1, 5)$

and its log transformation is

$$\log y = \log f(\boldsymbol{x}; \boldsymbol{\beta}) = \beta_0 + \beta_1 \log x_1 + \cdots + \beta_m \log(x_m) \equiv \boldsymbol{x}^\top \boldsymbol{\beta},$$

where $\boldsymbol{x} = (1, \log(x_1), \ldots, \log(x_m))^\top$ is a $(m+1)$-vector input in the natural log form.

With the data of n firms, the log-likelihood function for the Cobb–Douglas production given in Eq. (1) for MSN-SFM and SN-SFM is given as follows, respectively:

$$\begin{aligned} \ell_{MSN}(\boldsymbol{\beta}, \boldsymbol{\theta}) &= \sum_{i=1}^{n} \log f_{\mathcal{E}}(\log y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}) \\ &= logC + \log \phi_n \left(\boldsymbol{\varepsilon}; \boldsymbol{0}_n, \boldsymbol{\Sigma} + \sigma_u^2 \boldsymbol{I}_n \right) + \log \Phi_{n+1} \left(\boldsymbol{D}\boldsymbol{\varepsilon}; \boldsymbol{0}_{n+1}, \boldsymbol{\Delta} \right), \end{aligned}$$

and

$$\ell_{SN}(\boldsymbol{\beta},\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f_{\mathcal{E}}(\log y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}) = n \log C + \sum_{i=1}^{n} \log \phi(\log y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}; 0, \sigma_v^2 + \sigma_u^2)$$
$$+ \sum_{i=1}^{n} \log \Phi_2(D(\log y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}); (0,0)^\top, \boldsymbol{\Delta}),$$

where $\boldsymbol{\theta} = (\sigma_v, \sigma_u, \lambda)$, y_i is the output from ith firm, $\boldsymbol{x}_i$ is the input vector for ith firm in log form and $f_{\mathcal{E}}(\cdot;\boldsymbol{\theta})$ is the density function given in Eq. (3).

In this section, the maximum likelihood (ML) estimation is implemented via the Nelder–Mead algorithm in the R package `maxLik` [19]. The new proposed multivariate skew normal-based SFMs have multidimensional parameters and may be multimodal, and the initial values are needed. For the initial values of $\boldsymbol{\beta}$, σ_v, and σ_u, we propose to use the maximum likelihood estimator (MLE) of classical SFM from the `frontier` package [20] in R [21]. Once the initial values of $\boldsymbol{\beta}$ are obtained, we can calculate the residuals $\boldsymbol{\varepsilon}$'s with $\varepsilon_i = y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}$, for $i = 1, \ldots, n$. For the initial value of parameter λ, we first compute the unbiased moment estimators for skewness in ε_i's, named the adjusted Fisher–Pearson standardized moment coefficient $\hat{\gamma} = \frac{\hat{\gamma_0} n^2}{(n-1)(n-2)}$, where $\hat{\gamma_0}$ is the sample skewness of ε_i's [22]. And we solve the initial value of λ by setting $\hat{\gamma}$ equal to the skewness formula given in Eq. (4).

Noted that the maximum likelihood estimators (MLE), $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}})$ of $(\boldsymbol{\beta}, \boldsymbol{\theta})$ in the SN-SFM and MSN-SFM can be used in TE' by

$$\hat{TE}(\hat{\varepsilon}_i; \hat{\boldsymbol{\theta}}) = \int_0^\infty \exp(-u) f_{U|\mathcal{E}}(u|\hat{\varepsilon}_i; \hat{\boldsymbol{\theta}}) du,$$

where $\hat{\varepsilon}_i = \log y_i - \boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}}$, and $f_{U|\mathcal{E}}(u|\varepsilon;\boldsymbol{\theta})$ is given in Eq. (5).

Next, the performance of the proposed MSN-SFM is assessed with simulation studies. To simulate the datasets, the SFM with an explanatory variable X is utilized:

$$\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{V} - \boldsymbol{U},$$

where $\boldsymbol{Y}$ represents the output vector, $X = \begin{pmatrix} 1, & \ldots, & 1 \\ x_1, & \ldots, & x_n \end{pmatrix}^\top$, with x_i's simulated from uniform (0, 1), and $\boldsymbol{\beta} = (0, \beta_1)^\top$ the vector of regression coefficients. $\boldsymbol{U} \sim \mathcal{HN}_n(\boldsymbol{0}, \sigma_u^2 I_n)$ is the vector of inefficiency error terms. $\boldsymbol{V}$ denotes the vector of measurement errors so that V_i's follow: (a) the i.i.d. $\mathcal{N}_1(0, \sigma_v^2)$ for classical SFM, (b) the i.i.d. $\mathcal{SN}_1(0, \sigma_v^2, \lambda)$ for SN-SFM and (c) the joint multivariate skew normal distribution $\mathcal{SN}_n(0, \boldsymbol{\Sigma}, \alpha)$ for MSN-SFM, where Σ and α are given in Corollary 3.1. In order to show the performance of our main results, we simulate data from MSN-SFM with two simulation scenarios in the simulation study. Under each simulation scenario, we simulated $M = 1000$ datasets with sample sizes $n = 25$, 50 and 100, respectively. Estimates of the parameters in standard SFM, SN-SFM and MSN-SFM are listed in Tables 1 and 2.

Akaike information criterion (AIC) and Bayesian information criterion (BIC) are the most commonly used model selection statistics for parametric models [23]. AIC and BIC are calculated respectively, by

$$\text{AIC} = 2\ \text{length}(\boldsymbol{\theta}) - 2\ell(\hat{\boldsymbol{\theta}}), \qquad \text{BIC} = \log(n)\ \text{length}(\boldsymbol{\theta}) - 2\ell(\hat{\boldsymbol{\theta}}),$$

where ℓ is the log-likelihood function; $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate of $\boldsymbol{\theta}$; length$(\boldsymbol{\theta})$ is the number of parameters in the model; and n is the sample size. Models are selected with minimize AIC or BIC values. Note that $\boldsymbol{\theta}_{SFM} = (\beta_1, \sigma_v, \sigma_u)$, $\boldsymbol{\theta}_{SN-SFM} = (\beta_1, \sigma_v, \sigma_u, \lambda)$, $\boldsymbol{\theta}_{MSN-SFM} = (\beta_1, \sigma_v, \sigma_u, \lambda)$, and both SN-SFM and MSN-SFM have the same number of parameters which have one more parameter than classical SFM.

Case 1 Data simulated from MSN-SFM: $\beta_1 = 1$, $\boldsymbol{U} \sim \mathcal{HN}_n(0, \sigma_u^2 \boldsymbol{I}_n)$ with $\sigma_u = 2$, $V_1, \ldots, V_n$ in $\boldsymbol{V}$ are identically distributed with $\boldsymbol{V} \sim \mathcal{SN}_n(0, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ with $\boldsymbol{\Sigma}$ and $\boldsymbol{\alpha}$ in Corollary 3.1 and $\sigma_v = 4$, $\lambda = -5$.

Table 1 shows the results with data simulated from MSN-SFM. The mean, absolute value of bias, and the standard deviation of the MLEs for the parameters in the SFM, SN-SFM and MSN-SFM are reported. We observe the following facts. (i) All biases of the MLEs for β_1, σ_u and σ_v in the SN-SFM and MSN-SFM are smaller than those in the SFM for three sample sizes. (ii) As the sample size n increases from 25 to 100, the bias in β_1, σ_u and σ_v in all three models SFM, SN-SFM and MSN-SFM decreases. (iii) AIC and BIC values for MSN-SFM are lower than those in both SN-SFM and SFM for all $n = 25$, 50 and 100.

Case 2 Data simulated from MSN-SFM with "wrong skewness": $\beta_1 = 1$, $\boldsymbol{U} \sim \mathcal{HN}_n(0, \sigma_u^2 \boldsymbol{I}_n)$ with $\sigma_u = 2$, $V_1, \ldots, V_n$ in $\boldsymbol{V}$ are identically distributed with $\boldsymbol{V} \sim \mathcal{SN}_n(0, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ with $\boldsymbol{\Sigma}$ and $\boldsymbol{\alpha}$ in Corollary 3.1 and $\sigma_v = 4$, $\lambda = 5$.

Table 2 shows simulation results with data simulated from MSN-SFM. Note that the simulated data show the positive skewness, where the negative skewness is expected in SFM. Thus, the SFM fails to provide estimates for this simulation scenario and the mean, absolute value of bias and the standard deviation of MLEs for the parameters in the SN-SFM and MSN-SFM are reported. We observe the following facts. (i) All biases of the MLEs for β_1, σ_u, σ_v and λ in the MSN-SFM are smaller than those in SN-SFM for three sample sizes. (ii) As the sample size n increases from 25 to 100, both the bias and the standard deviation for β_1, σ_u, σ_v and λ in MSN-SFM decrease. (iii) AIC and BIC values for MSN-SFM are lower than those in both SN-SFM for all $n = 25$, 50 and 100.

6 Real Data Analysis

In this section, the real data are from rice producers in the Tarlac region of the Philippines in the year 1992, [24, 25]. The dataset from the first $n = 20$ rice producers presents a positive skewness for residuals under the classical normal–half-normal

Table 1 The sample mean, absolute value of bias and variance of the MLE for parameters in the SFM, SN-SFM and MSN-SFM are given for sample sizes $n = 25$, 50 and 100 for $M = 1000$ simulated datasets

$n = 25$	SFM			SN-SFM				MSN-SFM (true model)			
	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	−5	1	4	2	−5
Mean	−2.592	0.379	3.365	0.383	1.706	1.060	2.450	0.689	3.180	1.848	−4.799
| Bias |	3.592	3.621	1.365	0.617	2.294	0.940	7.450	0.311	0.820	0.152	0.201
Std.Dev.	3.023	0.522	1.404	0.937	0.585	0.800	1.866	0.952	2.606	0.635	6.492
Log−L.	−49.910			−47.584				−43.997			
AIC	105.819			**103.169**				**95.995**			
BIC	109.476			**108.044**				**100.870**			
$n = 50$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	−5	1	4	2	−5
Mean	−2.306	0.398	3.339	0.547	1.763	1.103	3.022	0.796	3.171	1.922	−4.239
|Bias|	3.306	3.602	1.339	0.453	2.237	0.897	8.022	0.204	0.829	0.078	0.761
Std.Dev.	2.727	0.569	1.241	0.778	0.524	0.830	2.737	0.673	2.311	0.498	4.495
Log-L.	−100.373			−95.197				−88.097			
AIC	206.747			**198.394**				**184.194**			
BIC	212.483			**206.042**				**191.842**			
$n = 100$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	−5	1	4	2	−5

Table 1 continued

$n = 25$	SFM			SN-SFM				MSN-SFM (true model)			
	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
Mean	−2.676	0.376	3.627	0.548	1.880	1.180	3.366	0.963	3.737	1.965	−4.714
|Bias|	3.676	3.624	1.627	0.452	2.120	0.820	8.366	0.037	0.263	0.035	0.286
Std.Dev.	2.836	0.561	1.330	0.928	0.526	0.855	2.746	0.467	1.663	0.349	2.209
Log-L.	−207.278			−195.773				−176.929			
AIC	420.556			**399.545**				**361.858**			
BIC	428.372			**409.966**				**372.278**			

The log-likelihood (Log-L), AIC and BIC values are also shown. Bold-faced values indicate that the SN-SFM and MSN-SFM are preferable

Table 2 The sample mean, absolute value of bias and variance of the MLE for parameters in the SFM, SN-SFM and MSN-SFM are given for sample sizes $n = 25$, 50 and 100 for $M = 1000$ simulated datasets

$n = 25$	SFM			SN-SFM				MSN-SFM (true model)			
	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	5	1	4	2	5
Mean	x	x	x	1.469	1.548	0.225	2.697	1.052	3.069	1.180	5.015
|Bias|	x	x	x	0.469	2.452	1.775	2.303	0.052	0.931	0.820	0.015
Std.Dev.	x	x	x	1.220	0.488	1.421	4.783	1.006	2.019	1.462	5.583
Log-L.	x			−44.632				−43.883			
AIC	x			97.264				95.766			
BIC	x			102.139				100.642			
$n = 50$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	5	1	4	2	5
Mean	x	x	x	1.542	1.566	0.267	2.556	1.025	3.752	1.690	5.523
|Bias|	x	x	x	0.542	2.434	1.733	2.444	0.025	0.248	0.310	0.523
Std.Dev.	x	x	x	0.966	0.435	1.515	4.062	0.678	1.528	1.055	3.853
Log-L.	x			−90.510				−88.429			
AIC	x			189.020				184.858			
BIC	x			196.668				192.506			
$n = 100$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$

Table 2 continued

$n = 25$	SFM			SN-SFM				MSN-SFM (true model)			
	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$
True value	1	4	2	1	4	2	5	1	4	2	5
Mean	x	x	x	1.596	1.569	0.247	2.689	1.015	4.121	1.940	5.269
|Bias|	x	x	x	0.596	2.431	1.753	2.311	0.015	0.121	0.060	0.269
Std.Dev.	x	x	x	0.996	0.506	1.604	5.883	0.428	0.905	0.425	0.968
Log-L.	x			−182.347				−175.993			
AIC	x			372.695				359.986			
BIC	x			383.115				370.407			

The log-likelihood (Log-L), AIC and BIC values are also shown. SFM failed to work due to the "wrong" skewness

Table 3 Estimates of parameters from first $n = 20$ farms in Philippines rice production data in 1992 are given for the standard SFM, SN-SFM and MSN-SFM

SFM	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}_v$	$\hat{\sigma}_u$		Log-L.	AIC	BIC
Estimates	−.226	.714	.263	.054	.182	.0007		5.669	.661	−5.347
SN-SFM	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	Log-L.	AIC	BIC
Estimates	−.163	.771	.067	.249	.184	**.235**	6.872	6.216	1.566	**−6.442**
MSN-SFM	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}_v$	$\hat{\sigma}_u$	$\hat{\lambda}$	Log-L.	AIC	BIC
Estimates	1.456	.967	-.034	.031	.331	**1.81**	7.422	7.196	**−.392**	**−8.4**

The log-likelihood (Log-L), AIC and BIC values are also shown
Bold-faced values indicate that the SN-SFM and MSN-SFM are preferable

SFM. The output variable is the freshly threshed rice in tonnes (denoted by Y); the input variables are area planted in hectares (denoted by X_1), labor used including man-days of both family and hired labor (denoted by X_2) and fertilizer used in kilogram of active ingredients (denoted by X_3).

We apply the standard SFM, SN-SFM and MSN-SFM to the dataset. The Cobb–Douglas production function is applied as considered in [24, 25]:

$$\log(\mathrm{Y}_i) = \beta_0 + \beta_1 \log(\mathrm{X}_{1i}) + \beta_2 \log(\mathrm{X}_{2i}) + \beta_3 \log(\mathrm{X}_{3i}) + V_i - U_i, \qquad (7)$$

where U_i's are assumed to follow the i.i.d. $\mathcal{HN}_1(0, \sigma_u^2)$. The model Eq. (7) is called

(a) the SFM if V_i's are assumed to follow the i.i.d. normal $\mathcal{N}_1(0, \sigma_v^2)$ [26];
(b) the SN-SFM if V_i's are assumed to follow i.i.d. skew normal distribution, $\mathcal{SN}_1(0, \sigma_v^2, \lambda)$ given in Theorem 3.2;
(c) the MSN-SFM if V_i's follow the joint multivariate skew normal, $\mathcal{SN}_n(\mathbf{0}, \mathbf{\Sigma}, \alpha)$ given in Corollary 3.1.

From the maximum likelihood estimation results given in Table 3, we have the following conclusions: (i) The residuals from SFM estimation show a positive skewness (0.38), which is assumed to be negative under the classical SFM. The wrong skewness problem is evident under the classical SFM, which implies all farms in the industry are fully efficient. Indeed, there is no variability in inefficiency term U under SFM, where the MLE of σ_u is close to zero ($\hat{\sigma}_u < 0.001$). (ii) The SN-SFM and MSN-SFM provide strong positive skewness parameter estimates $\hat{\lambda} = 6.872$ and 7.422, respectively. (iii) The AIC values (0.661 for SFM, 1.566 for SN-SFM and −0.392 for MSN-SFM) indicate that MSN-SFM fits best among three models. (iv) The BIC values (−5.347 for SFM, −6.442 for SN-SFM and −8.4 for MSN-SFM) show that the MSN-SFM is preferred to the SN-SFM, and the SN-SFM is in turn preferred to the SFM.

Table 4 provides the descriptive statistics of the technical efficiency scores among farms (left panel), five farms with highest efficiency scores under SN-SFM and MSN-SFM (middle panel) and five farms with lowest efficiency scores under SN–SFM and MSN-SFM (right panel). Note that the skewness for residuals of the production function is positive so that the classical SFM gives close to zero estimate in variance of the inefficiency term U (Table 3). This results full efficiencies estimated for all

Table 4 The descriptive statistics of technical efficiency for Philippines rice production data in 1992 for the standard SFM, SN-SFM and MSN-SFM are given in left panel

				Highest five farm code			Least five farm code		
	SFM	SN-SFM	MSN-SFM	SFM	SN-SFM	MSN-SFM	SFM	SN-SFM	MSN-SFM
Mean	1	.836	.637	x	4	7	x	11	14
S.D.	0	.094	.128	x	(.946)	(.778)	x	(.635)	(.411)
Min	1	.635	.411	x	12	20	x	8	11
Max	1	.946	.778	x	(.917)	(.776)	x	(.675)	(.416)
				x	18	12	x	3	3
				x	(.907)	(.775)	x	(.69)	(.471)
				x	7	4	x	14	15
				x	(.906)	(.775)	x	(.708)	(.49)
				x	20	17	x	15	8
				x	(.902)	(.77)	x	(.738)	(.5)

Five farms with the highest efficiency scores and five farms with the lowest efficiency scores are shown in the middle panel and right panel, respectively. Efficiency scores are given in parentheses. The standard SFM fails to give the ranks among farms due to the wrong skewness

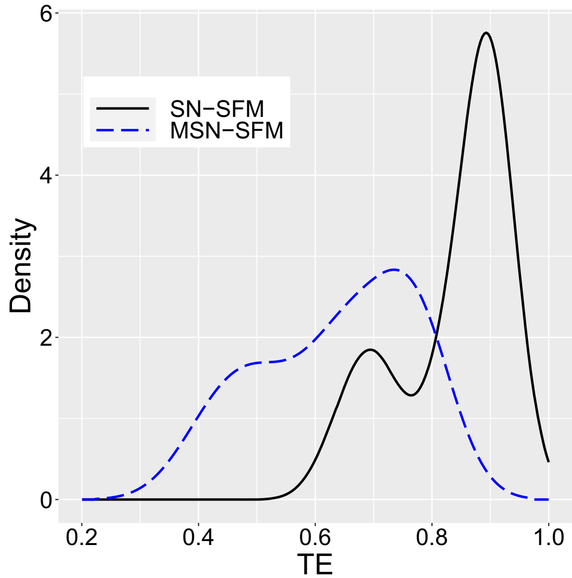

Fig. 4 The kernel density of efficiency scores for SN-SFM and MSN-SFM using Philipp -ines rice production data in 1992. The standard SFM fails to generate the TE scores among farms due to the wrong skewness

farms and the classical SFM fails to work in this case. Furthermore, Fig. 4 shows the kernel densities for empirical distributions of efficiency scores under SN-SFM (blue dashed curve) and MSN-SFM (black solid curve), respectively. From middle and right panel of Table 4 and Fig. 4, we can see that the TE scores under SN-SFM are higher than those given by MSN-SFM, and the TE ranks under SN-SFM and MSN-SFM are similar for most of 20 farms. For example, farms with the highest five TE scores and with the least five TE scores are the same farms under both proposed models. The new proposed SN-SFM and MSN-SFM in Sect. 3 provide nonzero TE (Table 4) and thus have the ability to estimate the variability in the industry and solved the wrong skewness problem.

7 Conclusion and Future Works

In this paper, we proposed a new stochastic frontier model based on multivariate skew normal distribution. The proposed model assumes the noise term V follows the multivariate skew normal distribution. In SFM literature, the classical SFM assumes i.i.d. normal distributions in the noise term. First, the i.i.d. assumption simplifies the underlying formulas and estimation. However, the i.i.d. assumption may or may not always be realistic in applications and the estimators under the dependence assumption may have smaller mean square error than those under the independent assumption [11]. Second, it is known that by assuming normality in classical SFM, the composite error term always has the negative skewness. Our multivariate skew normal-based SFM provides a solution to the "wrong" skewness problem.

The proposed model requires evaluations of Φ_n, and it may be difficult to calculate for large sample size. As one reviewer pointed out, the derivation of expectation–

maximization (EM) algorithm using the stochastic representation of multivariate skew normal distribution for parameter estimation is an important and interesting topic to investigate and may give the closest forms of the ML estimators in MSN-SFM. To extend the EM algorithm under current setting, one may need to investigate moments of the multivariate truncated closed skew normal distributions. A valuable future work would be a thorough investigation of developing EM algorithm for skew normal-based SFM under a feasible stochastic representation of multivariate skew normal distributions [27, 28].

Mathematica [29] codes for Fig. 2 and R [21] codes for the rest of figures, simulations and real data analysis are available upon the request.

Acknowledgements The authors would like to thank reviewers and Professor Hung T. Nguyen for their valuable suggestions and comments, which improved the manuscript. The first author would like to thank the support of Elizabeth Gaines Mann Professorship from University of North Alabama.

Funding This work was not supported by any funding source.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Aigner D, Lovell CK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econ 6(1):21–37
2. Meeusen W, van Den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Int Econ Rev 18:435–444
3. Andrade BB, Souza GS (2019) The EM algorithm for standard stochastic frontier models. Pesqui Oper 39:361–378
4. Greene WH (1990) A gamma-distributed stochastic frontier model. Journal of econometrics 46(1–2):141–163
5. Green A, Mayes D (1991) Technical inefficiency in manufacturing industries. Econ J 101(406):523–538
6. Wei Z, Zhu X, Wang T (2021) The extended skew-normal-based stochastic frontier model with a solution to 'wrong skewness' problem. Statistics 3:1–20
7. Smith MD (2008) Stochastic frontier models with dependent error components. Economet J 11(1):172–192
8. Bonanno G, De Giovanni D, Domma F (2017) The wrong skewness problem: a re-specification of stochastic frontiers. J Prod Anal 47(1):49–64
9. Wei Z, Conlon E, Wang T (2020) Asymmetric dependence in the stochastic frontier model using skew normal copula. Int J Approx Reason 128:56–68
10. Azzalini A, Dalla Valle A (1996) The multivariate skew-normal distribution. Biometrika 83(4):715–726
11. Wang C, Wang T, Trafimow D, Talordphop K (2021) Estimating the location parameter under skew normal settings: Is violating the independence assumption good or bad?. Soft Comput 25:7795–7802
12. Azzalini A, Capitanio A (1999) Statistical applications of the multivariate skew normal distribution. J R Stat Soc Ser B (Stat Methodol) 61(3):579–602
13. Wang T, Li B, Gupta AK (2009) Distribution of quadratic forms under skew normal settings. J Multivar Anal 100(3):533–545
14. Ye R, Wang T, Gupta AK (2014) Distribution of matrix quadratic forms under skew-normal settings. J Multivar Anal 131:229–239
15. González-Farías G, Domínguez-Molina A, Gupta AK (2004) Additive properties of skew normal random vectors. J Stat Plan Inference 126(2):521–534

16. Colombi R, Kumbhakar SC, Martini G, Vittadini G (2014) Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. J Prod Anal 42(2):123–136
17. Battese GE, Coelli TJ (1988) Prediction of firm level technical inefficiencies with a generalized frontier production function. J Econ 38:387–399
18. Greene WH (2008) Econometric analysis, 6th edn. Prentice Hall, Upper Saddle River
19. Henningsen A, Toomet O (2011) maxlik: A package for maximum likelihood estimation in R. Comput Stat 26(3):443–458
20. Coelli T, Henningsen A, Henningsen MA (2020) Package 'frontier'. Technical report
21. R Core Team (2013) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. R Foundation for Statistical Computing. http://www.R-project.org/
22. Joanes DN, Gill CA (1998) Comparing measures of sample skewness and kurtosis. J R Stat Soc Ser D (The Stat) 47(1):183–189
23. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723
24. Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE (2005) An introduction to efficiency and productivity analysis. Springer, New York
25. Henningsen A (2014) Introduction to econometric production analysis with R. Department of Food and Resource Economics, University of Copenhagen
26. Adcock C, Azzalini A (2020) A selective overview of skew-elliptical and related distributions and of their applications. Symmetry 12(1):118
27. Arellano-Valle RB, Genton MG (2005) On fundamental skew distributions. J Multivar Anal 96(1):93–116
28. Arellano-Valle RB, Azzalini A (2006) On the unification of families of Skew-normal distributions. Scand J Stat 33(3):561–574
29. Wolfram Research, Inc.: Mathematica 11.0. https://www.wolfram.com