Statistics
A Journal of Theoretical and Applied Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gsta20

# The extended skew-normal-based stochastic frontier model with a solution to 'wrong skewness' problem

Zheng Wei, Xiaonan Zhu & Tonghui Wang

**To cite this article:** Zheng Wei, Xiaonan Zhu & Tonghui Wang (2021): The extended skew-normal-based stochastic frontier model with a solution to 'wrong skewness' problem, Statistics, DOI: 10.1080/02331888.2021.2004142

**To link to this article:** https://doi.org/10.1080/02331888.2021.2004142

Published online: 29 Nov 2021.

STATISTICS
https://doi.org/10.1080/02331888.2021.2004142

Taylor & Francis
Taylor & Francis Group

Check for updates

# The extended skew-normal-based stochastic frontier model with a solution to 'wrong skewness' problem

Zheng Wei[a], Xiaonan Zhu[b] and Tonghui Wang[c]

[a]Department of Mathematics and Statistics, University of Maine, Orono, ME, USA; [b]Department of Mathematics, University of North Alabama, Florence, AL, USA; [c]Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM, USA

**ABSTRACT**
In this paper, a new extended skew-normal-based stochastic frontier model (SFM) is proposed to investigate the relationship between the disturbance $U$ (representing technical inefficiency) and $V$ (representing noise). Two methods for deriving the distribution of the error, $\mathcal{E} = V - U$, in the model are studied, together with the discussion of the model with wrong skewness problems in detail. The conditional distribution of $U$ given $\mathcal{E} = \varepsilon$ and technical efficiency of the model is obtained. For the illustration of our results, simulation studies on the performance of the proposed two new models are given with comparisons of the classical normal-half normal SFM. Two real data examples, including the one with the wrong skewness under the classic SFM, show that our skew normal based SFMs perform better than the classic SFM. More importantly, the extended skew-normal based SFMs provide a non-zero inefficiency estimates among firms with the correction of the wrong skewness problem.

**ARTICLE HISTORY**
Received 17 March 2021
Accepted 5 October 2021

**KEYWORDS**
Extended skew-normal distribution; stochastic frontier model; technical efficiency; wrong skewness problem

## 1. Introduction

The stochastic frontier model (SFM) is a popular and useful tool throughout the frontier productivity modelling literature. The original formulation, proposed independently in [1,2], assumes the production frontier $h(\boldsymbol{x})$ describing the maximum output with the given input $\boldsymbol{x}$. But in reality, a producer or a firm often produces the outputs $y$ less than it could feasibly produce, i.e., the output function would lie below the frontier. Thus, the real production $Y$ can be written as

$$Y = h(\boldsymbol{x}) \cdot TE,$$

where $TE$ is called the *technical efficiency* and $TE \leq 1$. The producer obtains its production frontier if and only if $TE = 1$, otherwise $TE < 1$. In canonical SFM, this model is often expressed in the following form:

$$Y_i = h(\boldsymbol{x}_i|\boldsymbol{\beta}) \exp(\varepsilon_i) \quad i = 1, \ldots, n,$$

**CONTACT** T. Wang twang@nmsu.edu

where $Y_i$ is the production (or the logarithm of the production) of the $i$th firm, $\boldsymbol{x}_i$ is a $k$-vector of input quantities (or the transformations) of the $i$th firm, $\boldsymbol{\beta}$ is a vector of unknown parameters, $\mathcal{E}_i = V_i - U_i$, and $U_i$ are non-negative (one-sided) random variables with cumulative distribution function (CDF) $F_U(u)$ that accounts for the possible inefficiency of firms and $V_i$ are the random variables with CDF $F_V(v)$ representing the possible measurement errors (i.e., randomness or statistical noise), and $U_i$ and $V_i$ are assumed to be independent random variables [3,4]. The inefficiency of the firm $i$ can be measured as

$$\exp(-U_i) = \frac{Y_i}{h(\boldsymbol{x}_i|\boldsymbol{\beta})\exp(V_i)}, \quad i = 1, \dots, n.$$

Let $V_i \sim N(0, \sigma_v^2)$, the normal distribution with mean 0 and variance $\sigma_v^2$, $U_i \sim HN(0, \sigma_u^2)$, the half-normal distribution with scale parameter $\sigma_u^2$, and $V_i$s and $U_i$s are independent. And if $h(\boldsymbol{x}|\boldsymbol{\beta})$ is assumed to be Cobb–Douglas production function, i.e., $h(\boldsymbol{x}|\boldsymbol{\beta}) = \exp(\beta_0)x_1^{\beta_1}x_2^{\beta_2}\dots x_k^{\beta_k}$, where $\boldsymbol{\beta} = (\beta_0, \dots, \beta_k)^T$. Then, the stochastic frontier model can be written as [5]

$$\log Y_i = \log(\boldsymbol{x}_i)^T\boldsymbol{\beta} + V_i - U_i. \tag{1}$$

It follows that the probability density function (PDF) of $\mathcal{E}_i$ is

$$f_{\mathcal{E}}(\varepsilon) = \frac{2}{\sqrt{\sigma_v^2 + \sigma_u^2}}\phi\left(\frac{\varepsilon}{\sqrt{\sigma_v^2 + \sigma_u^2}}\right)\left[1 - \Phi\left(\frac{\frac{\sigma_u}{\sigma_v}\varepsilon}{\sqrt{\sigma_v^2 + \sigma_u^2}}\right)\right],$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of the univariate standard normal distribution. There is a close connection between the density of $\mathcal{E}_i$ and the skew-normal distribution [6,7].

It can be shown that the skewness for the composite error term $\mathcal{E}$ in Equation (1) is negative for the standard SFM. However, in many applications, the residuals could yield positive skewness, which is known as the wrong skewness phenomenon in SFM literature [8]. The wrong skewness problem implies the estimated variance of the inefficiency component $U$ is zero, which in turn indicates an empirical artefact that the inefficiency scores are zero for all firms across the whole industry [9–11].

To overcome the wrong skewness problem, Smith [12] first time in the literature challenges the independence assumption between $V$ and $U$ and employed the symmetric copula models in SFM. Bonanno et al. [10] applied the FGM copula to capture the dependence among $V$ following the generalized logistic distribution and $U$ following the exponential distribution. More recently, Wei et al. [9] further extended the copula-based SFM by the asymmetric copula, the skew-normal copula, and argued the wrong skewness could be due to the asymmetric dependence among $V$ and $U$. In this paper, we propose an alternative approach to deal with the wrong skewness problem by considering the asymmetric distribution assumption in $V$.

To this end, we construct the extended skew-normal distribution-based SFM and its nest model, the skew-normal-based SFM, named ESN-SFM and SN-SFM, respectively. It assumes that the random error term $V$ in SFM follows the extended skew-normal distribution, which provides a natural generalization for standard SFM. The skewness in the composite error $\mathcal{E}$ in the proposed model is derived, and it can be shown that the new proposed SFMs allow both positive and negative skewness. Thus, it provides an alternative model for the standard SFM for the wrong skewness problem. The new proposed model also provides an alternative model for copula-based SFMs consideblack in [9,10,12] with asymmetric distribution in $V$. For the illustration of our main results, simulation studies on the performance of the proposed extended skew-normal–half-normal SFM are given with comparisons of the classical normal–half-normal SFM and skew-normal–half-normal SFM. Two real data applications, including the one with the wrong skewness under the classic SFM, show that our (extended) skew-normal-based SFM performs better than classic SFM. More importantly, the (extended) skew-normal-based SFMs provide a non-zero inefficiency estimates among firms with the correction of the wrong skewness problem.

This paper is organized as follows. The notations and review of skew-normal families as preliminaries are given in Section 2. The main results on constructing the SN-SFM and ESN-SFM are provided in Section 3. In Section 3.3, the skewness of the new proposed SFMs are derived. The firm-specific efficiency is discussed in Section 3.4. The maximum-likelihood estimation of parameters together with the simulation study are included in Sections 4.1 and 4.2, respectively. In Section 5, two real data applications are obtained for the illustration of our main results.

## 2. Preliminaries

Throughout this paper, we use $M_{p\times m}$ to denote the set of all $p \times m$ matrices over the real field $\mathbb{R}$ and $\mathbb{R}^p = M_{p\times 1}$. For any $B \in M_{p\times m}$, $B^T$ is the transpose of $B$. $I_p \in M_{p\times p}$ is the identity matrix. $\phi_p(\cdot\,;\boldsymbol{\mu},\Sigma)$ and $\Phi_p(\cdot\,;\boldsymbol{\mu},\Sigma)$ are the probability density function (PDF) and cumulative distribution function (CDF), respectively, of a $p$-dimensional normal distribution with the mean vector $\boldsymbol{\mu} \in \mathbb{R}^p$ and covariance matrix $\Sigma$, simply, $\phi_p(\cdot\,;\Sigma)$ and $\Phi_p(\cdot\,;\Sigma)$ for the case when $\boldsymbol{\mu} = \mathbf{0}$.

The class of skew-normal distributions was introduced by Azzalini [13]. Its extensions and applications have been studied by many researchers [14–16]. A random variable $Z$ is said to have the extended standard skew-normal ($ESN$) distribution with the skewness parameter $\lambda \in \mathbb{R}$ and the shape parameter $\tau \in \mathbb{R}$, denoted as $Z \sim ESN(\lambda,\tau)$, if its PDF is given by

$$f(z;\lambda,\tau) = \frac{\phi(z)\Phi(\tau\sqrt{1+\lambda^2}+\lambda z)}{\Phi(\tau)}, \quad z \in \mathbb{R}.$$

Let $X = \mu + \sigma Z$. The distribution of $X$ is called the extended skew-normal distribution with the location parameter $\mu$, the scale parameter $\sigma^2$, the skewness parameter $\lambda$, and the shape parameter $\tau$, denoted by $X \sim ESN(\mu,\sigma^2,\lambda,\tau)$ [see Chapter 2 in 7]. The PDF of $X$ is given by

$$f_X(x;\mu,\sigma^2,\lambda,\tau) = \frac{1}{\Phi(\tau)\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\tau\sqrt{1+\lambda^2}+\lambda\,\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}. \tag{2}$$

**Remark 2.1:** Note that if $\tau = 0$, the distribution in Equation (2) is reduced to the skew-normal distribution with the location parameter $\mu$, the scale parameter $\sigma^2$, the skewness parameter $\lambda$, denoted by $X \sim SN(\mu,\sigma^2,\lambda)$. Also, if $\tau = 0$ and $\lambda \to \infty$, the density of $X$ in

Equation (2) is blackuced to the truncated normal distribution with pre-truncate mean $\mu$ and variance $\sigma^2$.

The following lemma will be used in deriving the skew-normal-based SFMs.

**Lemma 2.1:** *Let $Z_0 \sim N(0,1)$ and $Z_1 > 0$ be the truncated normal with pre-truncate mean $\mu$ and variance $\sigma^2$, and $Z_0$ and $Z_1$ are independent. The CDF of $X = Z_0 + \lambda Z_1$ for a constant $\lambda$ is given by*

$$F_X(x) = \frac{1}{\Phi(\mu/\sigma)} \int_{-\infty}^{x} \phi(z_0) \left[ \Phi\left( \frac{x - \lambda\mu - z_0}{\lambda\sigma} \right) - \Phi\left( \frac{-\mu}{\sigma} \right) \right] \mathrm{d}z_0.$$

***Proof:*** Let $Y = \lambda(Z_1 - \mu)$. Since the PDF of $Z_1$ is

$$f_{Z_1}(z_1) = \frac{\phi\left(\frac{z_1-\mu}{\sigma}\right)}{\sigma[1 - \Phi(-\mu/\sigma)]} = \frac{\phi\left(\frac{z_1-\mu}{\sigma}\right)}{\sigma\Phi(\mu/\sigma)}, \quad z_1 \geq 0,$$

the PDF of $Y$ is

$$f_Y(y) = \frac{\phi\left(\frac{y}{\lambda\sigma}\right)}{\lambda\sigma[1 - \Phi(-\mu/\sigma)]} = \frac{\phi\left(\frac{y}{\lambda\sigma}\right)}{\lambda\sigma\Phi(\mu/\sigma)}, \quad y \geq -\mu\lambda.$$

The CDF of $X' = Z_0 + Y$ is

$$\begin{aligned} F_{X'}(x') &= \int_{-\infty}^{x'} \int_{-\infty}^{\infty} \phi(z_0) f_Y(x - z_0) \mathrm{d}z_0 \, \mathrm{d}x \\ &= \frac{1}{\Phi(\mu/\sigma)} \int_{-\infty}^{x'+\mu\lambda} \phi(z_0) \left[ \Phi\left( \frac{x' - z_0}{\lambda\sigma} \right) - \Phi\left( \frac{-\mu}{\sigma} \right) \right] \mathrm{d}z_0. \end{aligned}$$

Thus, the CDF of $X = Z_0 + \lambda Z_1 = Z_0 + (\lambda Z_1 - \lambda\mu) + \lambda\mu = X' + \lambda\mu$ is

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X' + \lambda\mu \leq x) = F_{X'}(x - \lambda\mu) \\ &= \frac{1}{\Phi(\mu/\sigma)} \int_{-\infty}^{x} \phi(z_0) \left[ \Phi\left( \frac{x - \lambda\mu - z_0}{\lambda\sigma} \right) - \Phi\left( \frac{-\mu}{\sigma} \right) \right] \mathrm{d}z_0. \end{aligned}$$

■

The extended skew-normal distribution in Equation (2) is a special case of the closed skew-normal distribution. Recall that a random vector $\boldsymbol{X} \in \mathbb{R}^p$ is said to have the multivariate closed skew-normal distribution [17] with parameters $q \geq 1$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\Sigma \in M_{p\times p}$, $D \in M_{q\times p}$, $\boldsymbol{\nu} \in \mathbb{R}^q$, $\Delta \in M_{q\times q}$, denoted by $\boldsymbol{X} \sim CSN_{p,q}(\boldsymbol{\mu}, \Sigma, D, \boldsymbol{\nu}, \Delta)$, if its PDF is given by

$$f_{p,q}(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma, D, \boldsymbol{\nu}, \Delta) = C\phi_p(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma)\Phi_q(D(\boldsymbol{x} - \boldsymbol{\mu}); \boldsymbol{\nu}, \Delta), \quad \boldsymbol{x} \in \mathbb{R}^p, \tag{3}$$

where $C^{-1} = \Phi_q(\boldsymbol{0}; \boldsymbol{\nu}, \Delta + D\Sigma D^T)$. Note that $X \sim ESN(\mu, \sigma^2, \lambda, \tau)$ is equivalent to $X \sim CSN_{1,1}(\mu, \sigma^2, \frac{\delta}{\sigma}, -\tau, 1 - \delta^2)$, where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

The moment generating function (MGF) of $\boldsymbol{X}$, which will be utilized for constructing the ESN-based SFM, is given in the following lemma.

**Lemma 2.2 ([17]):** *Let $\boldsymbol{X} \sim CSN_{p,q}(\boldsymbol{\mu}, \Sigma, D, \boldsymbol{\nu}, \Delta)$, then the MGF of $\boldsymbol{X}$ is*

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \frac{\Phi_q(D\Sigma\boldsymbol{t}; \boldsymbol{\nu}, \Delta + D\Sigma D^T)}{\Phi_q(\boldsymbol{0}; \boldsymbol{\nu}, \Delta + D\Sigma D^T)} \exp\left(\boldsymbol{t}^T\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^T\Sigma\boldsymbol{t}\right), \quad \boldsymbol{t} \in \mathbb{R}^p. \tag{4}$$

## 3. The new skew-normal-based SFMs

In this section, we are going to employ two approaches for constructing skew-normal-based SFMs. The corresponding skewness formula and the firm-specific efficiency are derived.

### 3.1. The skew-normal–half-normal-based SFM

Assume that $V \sim SN(0, \sigma_v^2, \lambda)$ and $U \sim HN(0, \sigma_u^2)$ are independent random variables. Then, the joint PDF of $(V, U)$ is

$$f_{(V,U)}(v, u) = 4\phi\left(v; 0, \sigma_v^2\right)\Phi\left(\frac{\lambda}{\sigma_v}v\right)\phi\left(u; 0, \sigma_u^2\right), \quad -\infty < v < \infty, \quad 0 \le u < \infty.$$

Let the composite error $\mathcal{E} = V - U$. The joint distribution of $(U, \mathcal{E})$ can be derived as follows:

$$\begin{aligned} f_{(U,\mathcal{E})}(u, \varepsilon) &= 4\phi(u + \varepsilon; 0, \sigma_v^2)\Phi\left(\frac{\lambda}{\sigma_v}(u + \varepsilon)\right)\phi(u; 0, \sigma_u^2) \\ &= \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right) 4\phi\left(u; -\frac{\sigma_u^2\varepsilon}{\sigma_u^2 + \sigma_v^2}, \frac{\sigma_u^2\sigma_v^2}{\sigma_u^2 + \sigma_v^2}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u + \varepsilon)\right), \end{aligned}$$

for $0 \le u < \infty, -\infty < \varepsilon < \infty$. Let $\mu = -\frac{\sigma_u^2\varepsilon}{\sigma_u^2+\sigma_v^2}$ and $\sigma^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_u^2+\sigma_v^2}$. The joint density of $U$ and $\mathcal{E}$ can be rewritten as

$$\begin{aligned} f_{(U,\mathcal{E})}(u, \varepsilon) &= \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right)\frac{4}{\sigma}\phi\left(\frac{u - \mu}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u + \varepsilon)\right) \\ &= \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right) 4\left(1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right)\frac{\phi\left(\frac{u-\mu}{\sigma}\right)}{\sigma\left[1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right]}\Phi\left(\frac{\lambda}{\sigma_v}(u + \varepsilon)\right) \end{aligned} \tag{5}$$

for $0 \le u < \infty, -\infty < \varepsilon < \infty$. Note that the term $\frac{\phi(\frac{u-\mu}{\sigma})}{\sigma[1-\Phi(-\frac{\mu}{\sigma})]}$ in Equation (5) is the density of the truncated normal with pre-truncate mean $\mu$ and variance $\sigma^2$.

From Lemma 2.1 and the joint PDF of $U$ and $\mathcal{E}$ given in Equation (5), the marginal PDF of $\mathcal{E}$ can be derived as follows.

**Theorem 3.1:** *Assume that $V \sim SN(0, \sigma_v^2, \lambda)$ and $U \sim HN(0, \sigma_u^2)$ are independent random variables, and $\mathcal{E} = V - U$. Let $\mu = -\frac{\sigma_u^2\varepsilon}{\sigma_u^2+\sigma_v^2}$ and $\sigma^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_u^2+\sigma_v^2}$, then we have the following results.*

(a) *The PDF of $\mathcal{E}$ is*

$$f_{\mathcal{E}}(\varepsilon) = \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right) 4\left(1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right) E_U\left[\Phi\left(\frac{\lambda}{\sigma_v}(U + \varepsilon)\right)\right], \tag{6}$$

*for* $-\infty < \mathcal{E} < \infty$.

(b) *The conditional PDF of U given* $\mathcal{E} = \varepsilon$ *is*

$$f_{U|\mathcal{E}}(u|\varepsilon) = C\phi\left(\frac{u-\mu}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u+\varepsilon)\right), \tag{7}$$

*for* $-\infty < \mathcal{E} < \infty$, $0 \le u < \infty$, *where* $C^{-1} = \sigma(1 - \Phi(-\frac{\mu}{\sigma}))E_U[\Phi(\frac{\lambda}{\sigma_v}(U + \varepsilon))]$ *and*

$$E_U\left[\Phi\left(\frac{\lambda}{\sigma_v}(U+\varepsilon)\right)\right] = \frac{1}{\Phi(\mu/\sigma)}\int_{-\infty}^{\frac{\lambda}{\sigma_v}\varepsilon} \phi(z_0)\left[\Phi\left(\frac{\frac{\lambda}{\sigma_v}\varepsilon + \frac{\lambda}{\sigma_v}\mu - z_0}{-\frac{\lambda}{\sigma_v}\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)\right] dz_0.$$

***Proof:*** From the joint PDF of $U$ and $\mathcal{E}$ given in Equation (5), we obtain the marginal PDF of $\mathcal{E}$,

$$\begin{aligned} f_{\mathcal{E}}(\varepsilon) &= \int_0^\infty f(u,\varepsilon)\,du = \int_0^\infty \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right)\frac{4}{\sigma}\phi\left(\frac{u-\mu}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u+\varepsilon)\right) du \\ &= \phi\left(\varepsilon; 0, \sigma_u^2 + \sigma_v^2\right) 4\left(1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right) E_{Z_\varepsilon}\left[\Phi\left(\frac{\lambda}{\sigma_v}(Z_\varepsilon + \varepsilon)\right)\right], \end{aligned} \tag{8}$$

where $Z_\varepsilon$ follows the truncated normal with pre-truncate mean $\mu$ and variance $\sigma^2$. Furthermore, we can obtain the conditional PDF of $U$ given $\mathcal{E}$,

$$f_{U|\mathcal{E}}(u|\varepsilon) = \frac{f_{U,\mathcal{E}}(u,\varepsilon)}{f_{\mathcal{E}}(\varepsilon)} = C\phi\left(\frac{u-\mu}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u+\varepsilon)\right),$$

where $C^{-1} = \sigma(1 - \Phi(-\frac{\mu}{\sigma}))E_{Z_\varepsilon}[\Phi(\frac{\lambda}{\sigma_v}(Z_\varepsilon + \varepsilon))]$. From Lemma 2.1, we have

$$\begin{aligned} E_{Z_\varepsilon}\left[\Phi\left(\frac{\lambda}{\sigma_v}(Z_\varepsilon+\varepsilon)\right)\right] &= E_{Z_\varepsilon}\left[P\left(Z_0 \le \frac{\lambda}{\sigma_v}(Z_\varepsilon+\varepsilon)\right)\right] \\ &= E_{Z_\varepsilon}\left[P\left(Z_0 - \frac{\lambda}{\sigma_v}Z_\varepsilon \le \frac{\lambda}{\sigma_v}\varepsilon\right)\right] \\ &= \frac{1}{\Phi(\mu/\sigma)}\int_{-\infty}^{\frac{\lambda}{\sigma_v}\varepsilon}\phi(z_0) \\ &\quad \times \left[\Phi\left(\frac{\frac{\lambda}{\sigma_v}\varepsilon + \frac{\lambda}{\sigma_v}\mu - z_0}{-\frac{\lambda}{\sigma_v}\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)\right] dz_0, \end{aligned}$$

where $Z_0 \sim N(0,1)$ is independent of $Z_\varepsilon$ and $\mathcal{E}$. ∎

### 3.2. The extended skew-normal–half-normal-based SFM

In this subsection, the random error term $V$ is assumed to follow the extended skew-normal distribution. We further extend the skew-normal-based SFM proposed in the last subsection to the ESN-based SFM by the MGF method.

**Theorem 3.2:** *Assume $V \sim ESN(0, \sigma_v^2, \lambda, \tau)$, and $U \sim HN(0, \sigma_u^2)$. Then, the density of $\mathcal{E} = V - U$ is*

$$f_{\mathcal{E}}(\varepsilon) = C\phi\left(\varepsilon; 0, \sigma_\varepsilon^2\right)\Phi_2\left(D\varepsilon; (-\tau, 0)^T, \Delta\right), \tag{9}$$

*where $C^{-1} = \Phi_2(\mathbf{0}_2; (-\tau, 0)^T, \Delta + \sigma_\varepsilon^2 DD^T)$, $D = \frac{1}{\sigma_\varepsilon^2}(\delta\sigma_v, -\sigma_u)^T$, and*

$$\Delta = I_2 - \frac{1}{\sigma_\varepsilon^2}\begin{pmatrix} \delta^2\sigma_v^2 & -\delta\sigma_v\sigma_u \\ -\delta\sigma_v\sigma_u & \sigma_u^2 \end{pmatrix}, \quad \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}, \quad \sigma_\varepsilon^2 = \sigma_u^2 + \sigma_v^2.$$

***Proof:*** From Equation (2.40) in [7], the MGF of $V$ is

$$M_V(t) = \frac{\Phi(\tau + \delta\sigma_v t)}{\Phi(\tau)}\exp\left(\frac{1}{2}\sigma_v^2 t^2\right) = \frac{\Phi(\delta\sigma_v t; -\tau, 1)}{\Phi(0; -\tau, 1)}\exp\left(\frac{1}{2}\sigma_v^2 t^2\right).$$

Thus, $V$ also follows $CSN_{1,1}(0, \sigma_v^2, \frac{\delta}{\sigma_v}, -\tau, 1 - \delta^2)$. Furthermore, the MGF of $U$ is

$$M_U(t) = 2\Phi(\sigma_u t)\exp\left(\frac{1}{2}\sigma_u^2 t^2\right).$$

Then, the MGF of $\mathcal{E} = V - U$ is

$$M_\varepsilon(t) = E[e^{tV}e^{-tU}] = \frac{2\Phi(-\sigma_u t)\Phi(\delta\sigma_v t; -\tau, 1)}{\Phi(0; -\tau, 1)}\exp\left(\frac{1}{2}\sigma_\varepsilon^2 t^2\right).$$

Thus, $\mathcal{E}$ follows $CSN_{1,2}(0, \sigma_\varepsilon^2, D, (-\tau, 0)^T, \Delta)$, with $D = \frac{1}{\sigma_\varepsilon^2}(\delta\sigma_v, -\sigma_u)^T$,

$$\Delta = I_2 - \frac{1}{\sigma_\varepsilon^2}\begin{pmatrix} \delta^2\sigma_v^2 & -\delta\sigma_v\sigma_u \\ -\delta\sigma_v\sigma_u & \sigma_u^2 \end{pmatrix}.$$

Therefore, the density of $\mathcal{E}$ is $f_\varepsilon(\varepsilon) = C\phi(\varepsilon; 0, \sigma_\varepsilon^2)\Phi_2(D\varepsilon; (-\tau, 0)^T, \Delta)$, where $C^{-1} = \Phi_2(\mathbf{0}_2; (-\tau, 0)^T, \Delta + D\Sigma D^T)$. ■

The special case of Theorem 3.2 is given below and its proof is straightforward.

**Corollary 3.1:** *Assume $V \sim SN(0, \sigma_v^2, \lambda)$, and $U \sim HN(0, \sigma_u^2)$. Then, the density of $\mathcal{E} = V - U$ is*

$$f_{\mathcal{E}}(\varepsilon) = C_0\phi\left(\varepsilon; 0, \sigma_\varepsilon^2\right)\Phi_2\left(D\varepsilon; \mathbf{0}_2, \Delta\right), \tag{10}$$

*where $C_0^{-1} = \Phi_2(\mathbf{0}_2; \mathbf{0}_2, \Delta + \sigma_\varepsilon^2 DD^T)$, $\sigma_\varepsilon^2$, $D$, and $\Delta$ are given in Theorem 3.2.*

**Remark 3.1:** We can show that the PDF of $\mathcal{E}$ given in Equation (6) of Theorem 3.1 and that in Equation (10) of Corollary 3.1 are equivalent. This also implies

$$\int_0^\infty \frac{4}{\sigma}\phi\left(\frac{u-\mu}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma_v}(u+\varepsilon)\right)\,du \equiv \frac{\Phi_2(D\varepsilon; \mathbf{0}_2, \Delta)}{\Phi_2(\mathbf{0}_2; \mathbf{0}_2, \Delta + \sigma_\varepsilon^2 DD^T)}.$$

The density curves of $\mathcal{E}$ with various of $\lambda$ and $\tau$ are given in Figure 2.

### 3.3. The skewness of the composite error in SN-SFM and ESN-SFM

The classical SFM assumes the disturbances term $U$ and the random noise term $V$ are independent random variables following the half-normal distribution and the normal distribution, respectively. Under this normal–half-normal distribution model, we know that the one side error term $U$ has $E[U] = \sqrt{\frac{2}{\pi}}\sigma_u$, $Var[U] = (\frac{\pi-2}{\pi})\sigma_u^2$ and its skewness is $E[(U - E[U])^3] = -\sqrt{\frac{2}{\pi}}(\frac{\pi-4}{\pi})\sigma_u^3$, with $U$ positively skewed. Thus, the expectation and skewness of $\mathcal{E}$ can be obtained by $E[\mathcal{E}] = -\sqrt{\frac{2}{\pi}}\sigma_u$, and $E[(\mathcal{E} - E[\mathcal{E}])^3] = \sqrt{\frac{2}{\pi}}(\frac{\pi-4}{\pi})\sigma_u^3$, which is strictly negative. It is known that the classical SFM often suffers from the 'wrong skewness phenomenon' – an empirical artefact that the residuals of the production function may have a positive skewness, whereas a negative one is expected under the model, which leads to estimated full efficiencies of all firms [10,11,18].

In the following, we investigate the skewness of the skew-normal-based SFMs proposed in Section 3.1 and 3.2.

**Proposition 3.1:** *Let $V \sim ESN(0, \sigma_v^2, \lambda, \tau)$, and $U \sim HN(0, \sigma_u^2)$, then the skewness of $\mathcal{E}$ is given by*

$$E\left[\left(\frac{\mathcal{E} - E[\mathcal{E}]}{\sigma_\varepsilon}\right)^3\right] = \frac{E[\mathcal{E}^3] - 3E[\mathcal{E}]E[\mathcal{E}^2] + 2(E[\mathcal{E}])^3}{(E[\mathcal{E}^2] - (E[\mathcal{E}])^2)^3}, \tag{11}$$

*where the first three moments of $\mathcal{E}$ are obtained by taking the derivatives of the MGF of $\mathcal{E}$ and given below*

$$E(\mathcal{E}) = \frac{\delta\sigma_v\phi(\tau)}{\Phi(\tau)} - 2\phi(0)\sigma_u, \quad E(\mathcal{E}^2) = \sigma_\varepsilon^2 - \frac{4\delta\sigma_u\sigma_v\phi(0)\phi(\tau)}{\Phi(\tau)} + \frac{\delta^2\sigma_v^2\phi'(\tau)}{\Phi(\tau)},$$

$$E(\mathcal{E}^3) = \frac{3\delta\sigma_v\sigma_\varepsilon^2\phi(\tau)}{\Phi(\tau)} + \frac{\delta^3\sigma_v^3\phi''(\tau)}{\Phi(\tau)} - 6\sigma_u\sigma_\varepsilon^2\phi(0) - \frac{6\delta^2\sigma_u\sigma_v^2\phi(0)\phi'(\tau)}{\Phi(\tau)} - 2\sigma_u^3\phi''(0),$$

*where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, $\sigma_\varepsilon^2 = \sigma_v^2 + \sigma_u^2$, $\phi'$ and $\phi''$ are the first and second derivatives of $\phi$.*

**Remark 3.2:** It can be shown that the new proposed SFM models accommodate both negative and positive skewness. Specifically, the skewness of ESN-SFM in Equation (11) as a function of $\delta$ and $\tau$ when $\sigma_u = 2$, $\sigma_v = 4$ is given in Figure 1(a). The skewness is positive for some values $0.4 < \delta < 1$ and $-3.5 < \tau < 3.5$. The graph and contour plot of skewness on those values are given in Figure 1(b,c), respectively.

### 3.4. Firm-specific efficiency

In this section, we investigate the efficiency of individual firms based on the SFMs proposed in Sections 3.1 and 3.2. As the noise term $V$ and inefficiency term $U$ in SFM given in Equation (1) are latent variables, and only the composite error term $\mathcal{E}$ can be determined after parameters have been estimated, the technical efficiencies $TE = \exp(-U)$ are generally unknown. By following [12,19], the $TE$ can be measured by the following conditional

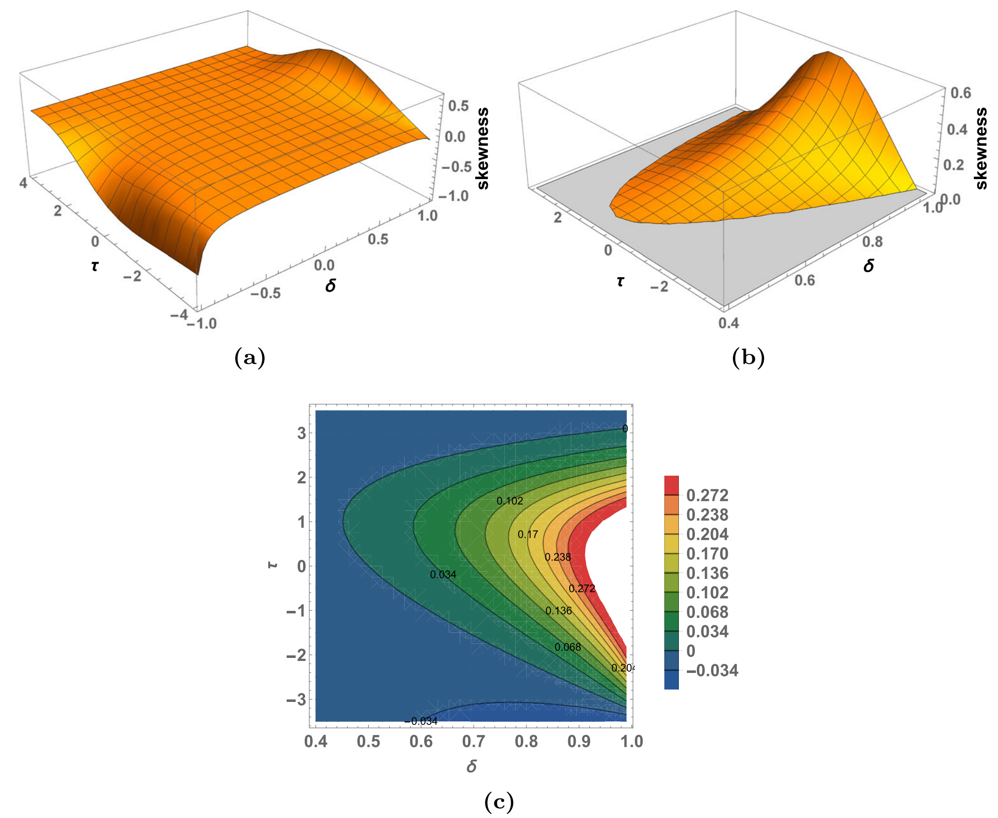


**Figure 1.** The skewness of $\mathcal{E}$ as a function of $\delta$ and $\tau$ for $\sigma_u = 2$ and $\sigma_v = 4$. Sub-figure (b) shows the range of positive skewness. Sub-figure (c) is the corresponding contour plot of sub-figure (b).

expectation:

$$TE(\varepsilon; \boldsymbol{\theta}) = E[\exp(-U)|\mathcal{E} = \varepsilon]. \tag{12}$$

Therefore, the conditional distribution of $U$ given $\mathcal{E} = \varepsilon$ is needed. From Theorem 3.2, we can obtain the conditional distribution of $U$ given $\mathcal{E} = \varepsilon$ and the TE, which is given below.

**Proposition 3.2:** *Let $V \sim ESN(0, \sigma_v^2, \lambda, \tau)$, and $U \sim HN(0, \sigma_u^2)$, then the conditional PDF of $U$ given $\mathcal{E} = \varepsilon$ is*

$$f_{U|\mathcal{E}}(u|\varepsilon) = \frac{2\phi_2\left(u+\varepsilon, u; (0,0)^T, diag(\sigma_v^2, \sigma_u^2)\right)\Phi\left(\tau\sqrt{1+\lambda^2} + \frac{\lambda}{\sigma_v}(u+\varepsilon)\right)}{C\phi(\varepsilon; 0, \sigma_v^2 + \sigma_u^2)\Phi_2(D\varepsilon; (-\tau, 0)^T, \Delta)}, \tag{13}$$

*where $C = \Phi(\tau)/\Phi_2(\mathbf{0}_2; (-\tau, 0)^T, \Delta + D\Sigma D^T)$ and $\Sigma$, $D$ are given in Equation (9), respectively.*

*And the TE, i.e., the conditional expectation of $e^{-U}$ given $\mathcal{E} = \varepsilon$, is*

$$TE(\varepsilon) = E_U[\exp(-U)|\mathcal{E} = \varepsilon] = \int_0^\infty \exp(-u) f_{U|\mathcal{E}}(u|\varepsilon)\, du. \tag{14}$$

**Remark 3.3:** The PDF curves, $f_{\mathcal{E}}(\varepsilon)$, of $\mathcal{E}$ in Equation (9) and their corresponding technical efficiencies, $TE(\varepsilon)$, in Proposition 3.2 are given Figure 2. The PDFs in Equation (9) with $\sigma_u = 2$ and $\sigma_v = 4$ are shown in the left column of Figure 2, and the plots of the $TE$ in Equation (14) with $\varepsilon = -0.1$ and $\sigma_v = 4$ are given in the right column. The plots in the top panel of Figure 2 shows that density and TE curves are affected by values of $\lambda = -5, -2, 5$ with the fixed $\tau = -1$. The plots in the middle panel of Figure 2 indicate that the density and TE curves are affected by values of $\tau = -1, 0, 1$ with the fixed $\lambda = -5$. The plots in the bottom panel of Figure 2 shows the density and TE curves having wrong skewness (positive skewness) with $\lambda = (1, 11)$ and $\tau = -1.5$.

From Figure 2, we conclude that the parameters of $\lambda$ and $\tau$ play important roles in the density curves of $f_{\mathcal{E}}(\varepsilon)$ and the shapes of $TE(\varepsilon)$ by the following observations: (i) the density curves of $f_{\mathcal{E}}(\varepsilon)$ show both left-skewed and right-skewed shape depending on the values in $\lambda$ and $\tau$; (ii) as $\lambda$ increases from negative values to positive values, the shape of the density curves of $f_{\mathcal{E}}(\varepsilon)$ changes from left-skewed to the right-skewed gradually, and the shape of the curves for $TE$ becomes more steeper as $\sigma_u$ increases; (iii) as $\tau$ changes from a positive value to negative and $\lambda$ is negative, the density curves of $f_{\mathcal{E}}(\varepsilon)$ skew more to the left and the $TE$ becomes more steeper as $\sigma_u$ increases.

## 4. Estimations and simulations

In this section, the maximum-likelihood estimations of parameters in proposed new SFMs are developed, together with simulation studies.

### 4.1. Maximum-likelihood estimation

In this section, we provide maximum-likelihood estimates for parameters in the proposed SN-SFM, ESN-SFM, together with estimates of the associated technical efficiencies.

We assume a priori that the efficient output is the production function of inputs with a specific functional form

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

for some unknown parameter vector $\boldsymbol{\beta}$. For example, the Cobb–Douglas production function has the form

$$y = f(\boldsymbol{x}) = f(\boldsymbol{x}; \boldsymbol{\beta}) = \exp(\beta_0) x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k}, \tag{15}$$

and its log-transformation is

$$\log y = \log f(\boldsymbol{x}; \boldsymbol{\beta}) = \beta_0 + \beta_1 \log x_1 + \cdots + \beta_k \log(x_k) \equiv \boldsymbol{x}^T \boldsymbol{\beta}, \tag{16}$$

where $\boldsymbol{x}$ is a $(k+1)$-vector with the inputs (including a constant term) in the log form.

When there is a cross-section of $n$ firms, the log-likelihood function for the Cobb–Douglas production given in Equation (1) is given by

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log f_{\mathcal{E}}(\log y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}) = n \log C$$

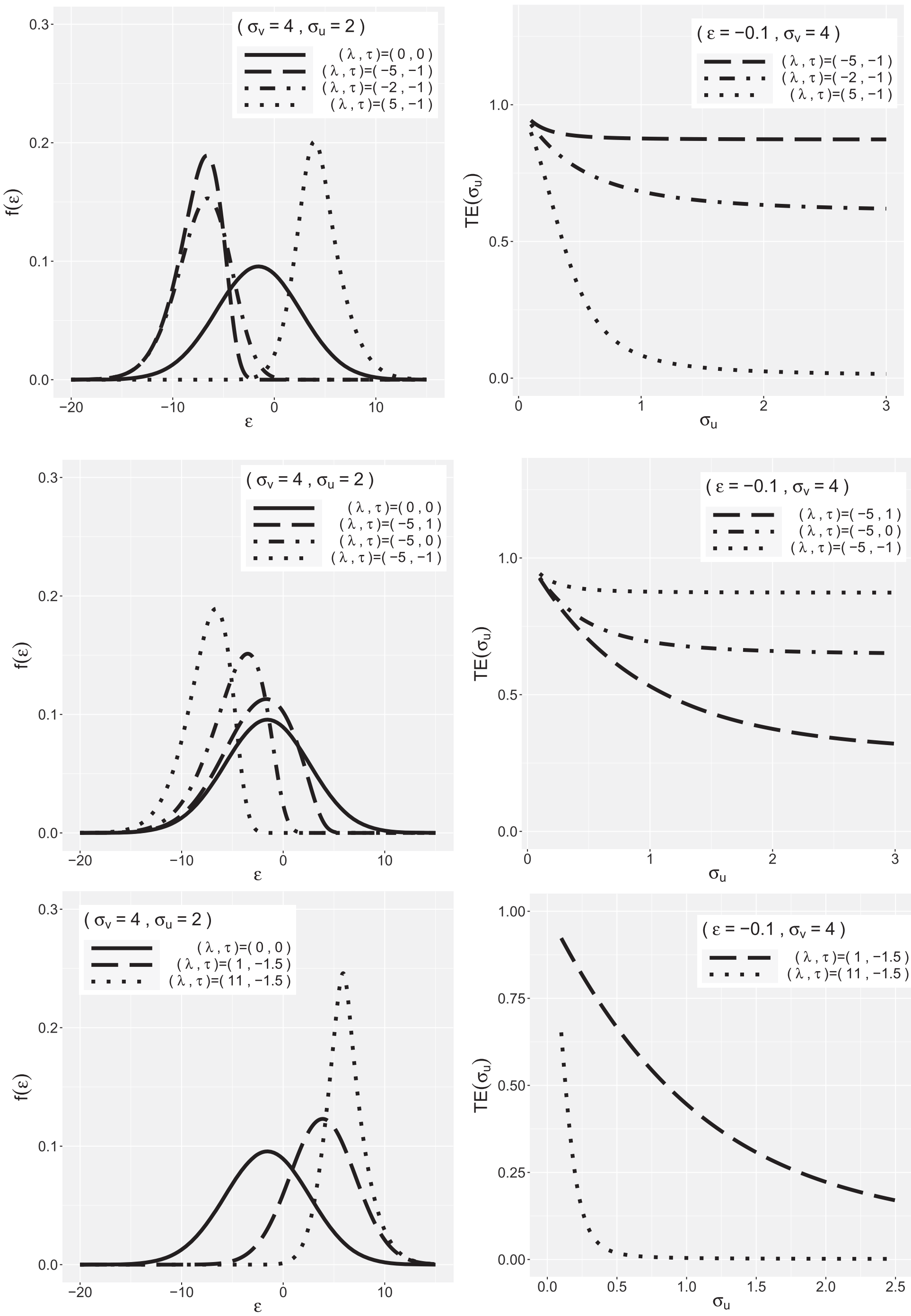


**Figure 2.** Density curves of $\varepsilon$ in Equation (9) with $\sigma_v = 4$ and $\sigma_u = 2$ given in left column, and curves of $TE(\varepsilon)$ in Equation (14) with $\varepsilon = -0.1$ and $\sigma_v = 4$ given in right column. The top panel plots are the curves with $\lambda = (-5, -2, 1)$ given fixed $\tau = -1$. The middle panel plots are the curves with $\tau = (-1, 0, 1)$ given fixed $\lambda = -5$. The bottom panel plots are the curves showing wrong skewness.

$$+\sum_{i=1}^{n}\log\phi(\log y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}; 0, \sigma_v^2+\sigma_u^2)$$
$$+\sum_{i=1}^{n}\log\Phi_2(D(\log y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}); (-\tau, 0)^T, \Delta), \tag{17}$$

where $\boldsymbol{\theta} = (\sigma_v, \sigma_u, \lambda, \tau)$, $y_i$ is the realization of the output from $i$th firm, $\boldsymbol{x}_i$ is the input vector for $i$th firm in log form, and $f_\varepsilon(\cdot; \boldsymbol{\theta})$ is the density function given in Equation (9).

In simulation studies and real data analysis, the maximum-likelihood (ML) estimation is implemented using the Nelder–Mead algorithm in the R package 'maxLik' [20], with the standard errors from the estimated variance–covariance matrix given in the final iteration. Since the vector of parameters in the new proposed skew-normal-based SFMs is multidimensional and may be multimodal, the initial values are needed. Here, we propose to use the initial values for $\boldsymbol{\beta}$, $\sigma_v$, and $\sigma_u$ as obtained from the maximum-likelihood estimator (MLE) of classic SFM from the 'frontier' package [21] in R [22]. From the initial value of $\boldsymbol{\beta}$, we can calculate the residuals $\varepsilon_i$'s with $\varepsilon_i = y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}$, for $i = 1, \dots, n$. For the initial value of parameter $\lambda$, we first compute the unbiased moment estimators for skewness in $\varepsilon_i$'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 $\varepsilon_i$'s [23]. And we solve the initial value of $\lambda$ by setting $\hat{\gamma}$ equal to the skewness formula given in Equation (11).

Once we obtain the MLEs $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}})$ of $(\boldsymbol{\beta}, \boldsymbol{\theta})$ in the SN-SFM and ESN-SFM, their associated TEs can be estimated by

$$\hat{TE}(\hat{\varepsilon}_i; \hat{\boldsymbol{\theta}}) = \int_0^\infty \exp(-u) f_{U|\mathcal{E}}(u|\hat{\varepsilon}_i; \hat{\boldsymbol{\theta}})\,\mathrm{d}u, \tag{18}$$

where $\hat{\varepsilon}_i = \log y_i - \boldsymbol{x}_i^T\hat{\boldsymbol{\beta}}$ and $f_{U|\mathcal{E}}(u|\varepsilon; \boldsymbol{\theta})$ is given in Equation (13).

We also compare models using the Akaike information criterion (AIC) which is the most commonly used model selection tool for parametric models [24]. AIC assigns the following values to the members of the parametric family,

$$\mathrm{AIC} = 2\,\mathrm{length}(\boldsymbol{\theta}) - 2\ell(\hat{\boldsymbol{\theta}}),$$

where $\ell$ is the log-likelihood function; $\hat{\boldsymbol{\theta}}$ is the maximum-likelihood estimate of $\boldsymbol{\theta}$; $\mathrm{length}(\boldsymbol{\theta})$ is the number of parameters in the model; and $n$ is the sample size. AIC is computed for each candidate model and models are selected with minimized AIC. AIC values are computed for each model in both simulation studies and real data analysis.

### 4.2. Simulation studies

In this section, the performance of the proposed SN-SFM, and the ESN-SFM, is assessed with simulation studies. To simulate the data sets, we consider the SFM

$$Y = \beta_1 X + V - U, \tag{19}$$

where $Y$ is the output function, $X$ is the single explanatory variable, distributed as $\mathrm{Uniform}(0, 1)$, and $\beta_1$ is the regression coefficient. $U$ is the inefficiency error term and

**Table 1.** The sample mean, absolute value of bias, and the standard deviation of the MLE for parameters in the SFM, SN-SFM, and ESN-SFM are given for sample sizes $n = 100$, 500 and 1000 under Case 1 for $M = 1000$ simulated data sets.

| | SFM | | | SN-SFM (true model) | | | | ESN-SFM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1 | 1 | 1 | 1 | 1 | 1 | −2 | 1 | 1 | 1 | −2 | 0 |
| Mean | 0.50 | 0.42 | 1.50 | 0.93 | 0.93 | 1.02 | −2.10 | 1.00 | 0.93 | 0.79 | −2.95 | −0.84 |
| \|Bias\| | 0.50 | 0.58 | 0.51 | 0.07 | 0.07 | 0.02 | 0.10 | 0.00 | 0.07 | 0.21 | 0.95 | 0.84 |
| Std. Dev. | 0.22 | 0.11 | 0.13 | 0.27 | 0.26 | 0.30 | 1.13 | 0.30 | 0.33 | 0.62 | 5.74 | 3.95 |
| $p_{AIC}$ | 0.13 | | | **0.74** | | | | 0.13 | | | | |
| $n = 500$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1 | 1 | 1 | 1 | 1 | 1 | −2 | 1 | 1 | 1 | −2 | 0 |
| Mean | 0.49 | 0.44 | 1.50 | 0.98 | 0.95 | 1.04 | −1.94 | 1.00 | 0.93 | 0.98 | −1.88 | −0.42 |
| \|Bias\| | 0.51 | 0.56 | 0.50 | 0.02 | 0.05 | 0.04 | 0.06 | 0.00 | 0.07 | 0.01 | 0.12 | 0.42 |
| Std. Dev. | 0.10 | 0.04 | 0.06 | 0.13 | 0.16 | 0.17 | 0.49 | 0.14 | 0.23 | 0.32 | 1.63 | 1.20 |
| $p_{AIC}$ | 0.00 | | | **0.85** | | | | 0.15 | | | | |
| $n = 1000$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1 | 1 | 1 | 1 | 1 | 1 | −2 | 1 | 1 | 1 | −2 | 0 |
| Mean | 0.49 | 0.44 | 1.50 | 0.98 | 0.94 | 1.05 | −1.90 | 1.00 | 0.92 | 1.03 | −1.74 | −0.35 |
| \|Bias\| | 0.51 | 0.56 | 0.50 | 0.02 | 0.06 | 0.05 | 0.10 | 0.00 | 0.08 | 0.03 | 0.26 | 0.35 |
| Std. Dev. | 0.07 | 0.03 | 0.04 | 0.08 | 0.12 | 0.12 | 0.37 | 0.09 | 0.19 | 0.21 | 0.74 | 0.74 |
| $p_{AIC}$ | 0.00 | | | **0.88** | | | | 0.12 | | | | |

Note: $p_{AIC}$ denotes the proportion of the models selected by using AIC among $M = 1000$ simulated data. Boldfaced values indicate that the SN-SFM and ESN-SFM are preferable.

follows the half-normal with the pre-truncation variance $\sigma_u^2$. $V$ denotes the measurement error term and follows $N(0, \sigma_v^2)$, $SN(0, \sigma_v^2, \lambda)$, and $ESN(0, \sigma_v^2, \lambda, \tau)$. In order to show the advantage of our main results, we use the following two sets of parameters for our simulation study. Under each simulation scenario, we simulated $M = 1000$ data sets with sample sizes $n = 100$, 500 and 1000, respectively. Estimates of the parameters in standard SFM, SN-SFM, and ESN-SFM under simulation scenarios 1 and 2 are listed in Tables 1 and 2, respectively.

*Case 1. Data simulated from SN-SFM*: $\beta_1 = 1$, $V \sim SN(0, \sigma_v^2, \lambda)$ with $\sigma_v = 1$, $\lambda = -2$, and $U \sim HN(0, \sigma_u^2)$ with $\sigma_u = 1$.

Table 1 shows the mean, absolute value of bias, and the standard deviation of the MLEs for the parameters in the SFM, SN-SFM and ESN-SFM for sample sizes $n = 100$, 500 and 1000, respectively. $p_{AIC}$ in Table 1 denotes the proportion of the model selected by using AIC among $M = 1000$ simulated data. We observe the following facts. (i) All biases of the MLEs for $\beta_1$, $\sigma_u$, and $\sigma_v$ in the SN-SFM and ESN-SFM are smaller than those in the SFM for three sample sizes. (ii) As the sample size $n$ increases, both the bias and standard deviations in $\beta_1$, $\sigma_u$, and $\sigma_v$ in all three models SFM, SN-SFM and ESN-SFM decrease. (iii) The likelihood values for SN-SFM and ESN-SFM are higher than those for SFM for all sample sizes. (iv) $p_{AIC}$ indicates the SN-SFM (true model) is the most frequently selected model by AIC among all three models with different sample sizes. Furthermore, as the sample size increases, $p_{AIC}$ for SN-SFM increases, while it decreases to zero for the SFM.

*Case 2. Data simulated from ESN-SFM*: $\beta_1 = 1$, $V \sim ESN(0, \sigma_v^2, \lambda, \tau)$ with $\sigma_v = 0.5$, $\lambda = -1$, and $\tau = -4$, and $U \sim HN(0, \sigma_u^2)$ with $\sigma_u = 0.5$.

Table 2 shows the mean, absolute value of bias, and the standard deviation of the MLEs for the parameters in the SFM, SN-SFM and ESN-SFM with sample sizes $n = 100$, 500 and

**Table 2.** The sample mean, absolute value of bias, and the standard deviation of the MLE for parameters in the SFM, SN-SFM, and ESN-SFM are given for sample sizes $n = 100$, 500 and 1000 under Case 2 for $M = 1000$ simulated data sets.

| | SFM | | | SN-SFM | | | | ESN-SFM (true model) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1.00 | 0.50 | 0.50 | 1.00 | 0.50 | 0.50 | −1.00 | 1.00 | 0.50 | 0.50 | −1.00 | −4.00 |
| Mean | −0.50 | 0.23 | 1.29 | 0.91 | 0.82 | 0.30 | 1.97 | 1.00 | 0.64 | 0.21 | −1.67 | −6.07 |
| \|Bias\| | 1.50 | 0.27 | 0.79 | 0.09 | 0.32 | 0.20 | 2.97 | 0.00 | 0.14 | 0.29 | 0.67 | 2.07 |
| Std.Dev. | 0.12 | 0.10 | 0.08 | 0.14 | 0.07 | 0.09 | 0.38 | 0.17 | 0.16 | 0.37 | 0.93 | 1.56 |
| $p_{AIC}$ | 0.00 | | | 0.57 | | | | 0.43 | | | | |
| $n = 500$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1.00 | 0.50 | 0.50 | 1.00 | 0.50 | 0.50 | −1.00 | 1.00 | 0.50 | 0.50 | −1.00 | −4.00 |
| Mean | −0.49 | 0.26 | 1.28 | 0.94 | 0.83 | 0.30 | 1.99 | 1.00 | 0.68 | 0.31 | −1.70 | −5.13 |
| \|Bias\| | 1.50 | 0.24 | 0.78 | 0.06 | 0.33 | 0.20 | 2.99 | 0.00 | 0.18 | 0.19 | 0.70 | 1.13 |
| Std.Dev. | 0.06 | 0.05 | 0.03 | 0.06 | 0.04 | 0.04 | 0.19 | 0.07 | 0.15 | 0.27 | 0.63 | 0.88 |
| $p_{AIC}$ | 0.00 | | | 0.28 | | | | **0.72** | | | | |
| $n = 1000$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\beta}_1$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\hat{\tau}$ |
| True value | 1.00 | 0.50 | 0.50 | 1.00 | 0.50 | 0.50 | −1.00 | 1.00 | 0.50 | 0.50 | −1.00 | −4.00 |
| Mean | −0.49 | 0.26 | 1.28 | 0.94 | 0.83 | 0.29 | 1.99 | 1.00 | 0.68 | 0.33 | −1.65 | −4.95 |
| \|Bias\| | 1.50 | 0.24 | 0.78 | 0.06 | 0.33 | 0.21 | 2.99 | 0.00 | 0.18 | 0.17 | 0.65 | 0.95 |
| Std.Dev. | 0.04 | 0.03 | 0.02 | 0.04 | 0.03 | 0.03 | 0.14 | 0.05 | 0.14 | 0.23 | 0.50 | 0.64 |
| $p_{AIC}$ | 0.00 | | | 0.13 | | | | **0.87** | | | | |

Note: $p_{AIC}$ denotes the proportion of the models selected by using AIC among $M = 1000$ simulated data.

1000, respectively. $p_{AIC}$ in Table 2 denotes the proportion of the model selected by using AIC among $M = 1000$ simulated data. We have the following conclusions. (i) All biases of the MLEs for $\beta_1$, $\sigma_u$, and $\sigma_v$ in the SN-SFM and ESN-SFM are smaller than those in the SFM for three sample sizes. (ii) As the sample size increases from $n = 100$ to 1000, the standard deviations in $\beta_1$, $\sigma_u$, and $\sigma_v$ in all three models SFM, SN-SFM and ESN-SFM decrease. (iii) The likelihood values for SN-SFM and ESN-SFM are higher than those for SFM for all sample sizes. (iv) $p_{AIC}$ indicates the ESN-SFM (true model) is the most frequently selected model by AIC among all three models for sample sizes $n = 500$ and $n = 1000$. Furthermore, as the sample size increases, $p_{AIC}$ for ESN-SFM increases.

## 5. Real data analysis

In this section, the new proposed stochastic frontier model is applied to two real data sets collected from $n = 43$ small holder rice producers in the Tarlac region of the Philippines, from the year 1992 and 1997 [25,26], respectively. The first data set is from the year 1992 since we find the presence of a strong positive skewness for residuals, while the negative skewness was expected from the classical normal–half-normal SFM. The second data set from the year 1997 does not evidence the wrong skewness and we investigate the impact of new proposed skew-normal-based SFM on the distribution of the efficiency scores.

The output variable is the freshly threshed rice in tonnes (denoted by Prod), the input variables include area planted in hectares (denoted by Area), labour used including man-days of both family and hiblack labour (denoted by Labour), fertilizer used in kilogram of active ingblackients (denoted by Npk).

We apply the standard SFM, SN-SFM and ESN-SFM to both data sets. For the underlying production function, the Cobb–Douglas production function is applied with the dependent variable Prod, and inputs variables Area, Labour and Npk as consideblack

**Table 3.** Estimates (with Standard Error in parentheses) of parameters from $n = 43$ farms in Philippines rice production data in 1992 are given for the standard SFM, SN-SFM and ESN-SFM. The log-likelihood (Log-L), AIC values are also shown.

| SFM | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | | | Log-L. | AIC |
|---|---|---|---|---|---|---|---|---|---|---|
| Estimates | −1.638 | .383 | .364 | .309 | .2 | .0001 | | | 8.257 | −4.514 |
| Standard Error | (1.519) | (.115) | (.137) | (.062) | (.022) | (1.813) | | | | |
| 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 |
| Estimates | −1.766 | .364 | .372 | .326 | .243 | .215 | 11.748 | | 9.502 | **−5.004** |
| Standard Error | (.371) | (.105) | (.118) | (.056) | (.047) | (.048) | (4.977) | | | |
| ESN-SFM | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\tau$ | Log-L. | AIC |
| Estimates | −2.376 | .337 | .387 | .331 | .373 | .214 | 10.729 | −1.5 | 9.824 | −3.648 |
| Standard Error | (.391) | (.091) | (.116) | (.057) | (.076) | (.044) | (3.462) | | | |

in [25,26]:

$$\log(Prod) = \beta_0 + \beta_1 \log(Area) + \beta_2 \log(Labour) + \beta_3 \log(Npk) + V - U,$$

where $U$ is assumed to follow the half-normal $HN(0, \sigma_u^2)$ and $V$ is assumed to follow the normal $N(0, \sigma_v^2)$ distribution in standard SFM [6], and the skew-normal distribution, $SN(0, \sigma_v^2, \lambda)$, in SN-SFM, and extended skew-normal, $ESN(0, \sigma_v^2, \lambda, \tau)$, in ESN-SFM, respectively. We reproduce the results given in [26,27] and compare these results to our proposed SN-SFM and ESN-SFM.

Maximum-likelihood (ML) estimation is implemented using the Nelder–Mead algorithm in the R package 'maxLik' [20], with the estimators' asymptotic variance–covariance matrix estimated at the final iteration. To fit the ESN-SFM for the data under consideration, we have estimated an ESN-SFM with $\tau = -1.5$ for both data sets. The value of $\tau$ was selected by the inspection of the histogram of the residuals and maximizing of MLE with fixed $\tau$ values over the range $\tau = 2, 1.5, 1, 0.5, -0.5, -1, -1.5, -2$. After maximization with respect to the other parameters, we obtained a maximized log-likelihood value with $\tau = -1.5$, which is treated as the fixed value.

### *5.1. Wrong skewness in the Philippines rice production data in 1992*

In this section, we fitted the standard SFM, SN-SFM, and ESN-SFM for rice production data in the Philippines in 1992 [25,26] and presented the estimation results in Table 3.

From the results given in Table 3, we have the following observations. (i) The residuals from ordinal least-square estimation show a positive skewness where the negative skewness is expected under the classical SFM. (ii) The wrong skewness problem is evident under the classical SFM which implies all farms in the industry are fully efficient, a fact that is falsified in most of the cases if one examines the relative productivities. Indeed, the variance of the inefficiency term is close to zero ($\hat{\sigma}_u = 0.0001$), there is no variability in inefficiency term $U$. (iii) The SN-SFM and ESN-SFM provide strong and significant positive skewness parameter estimates $\lambda = 11.748$ and 10.729. The new proposed SN-SFM and ESN-SFM in Section 3 provide non-zero TE (see Table 4) and thus have the ability to estimate the variability in the industry and solved the wrong skewness problem in these data. (iv) The AIC value ( $-5.004$ for SN-SFM) indicates SN-SFM is preferable to SFM. (v) We conclude from Table 4 and Figure 3 that the TE scores under SN-SFM and ESN-SFM are very close and the TE ranks under SN-SFM and ESN-SFM are the same for most of the 43 farms. For

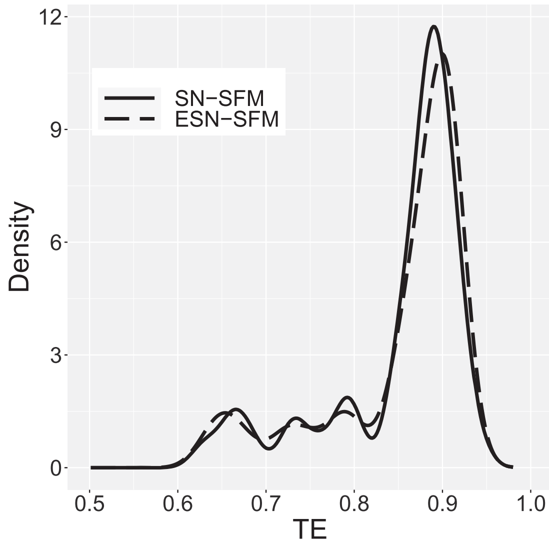


**Figure 3.** Kernel density of efficiency scores for standard SN-SFM and ESN-SFM using Philippines rice production data in 1992. The standard SFM fails to generate the TE scores among farms due to the wrong skewness.

**Table 4.** The descriptive statistics of Technical Efficiency for Philippines rice production data in 1992 for the standard SFM, SN-SFM and ESN-SFM are given in left panel. Five farms with highest efficiency scores and five farms with 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.

| | | | | Highest five farm code | | | Least five farm code | | |
|---|---|---|---|---|---|---|---|---|---|
| | SFM | SN-SFM | ESN-SFM | SFM | SN-SFM | ESN-SFM | SFM | SN-SFM | ESN-SFM |
| Mean | 1 | .851 | .852 | x | 4 | 4 | x | 34 | 34 |
| S.D. | 0 | .078 | .081 | x | (.938) | (.926) | x | (.631) | (.633) |
| Min | 1 | .631 | .633 | x | 31 | 31 | x | 8 | 8 |
| Max | 1 | .938 | .926 | x | (.921) | (.916) | x | (.660) | (.653) |
| | | | | x | 41 | 41 | x | 11 | 11 |
| | | | | x | (.917) | (.914) | x | (.665) | (.656) |
| | | | | x | 37 | 37 | x | 40 | 40 |
| | | | | x | (.915) | (.913) | x | (.682) | (.679) |
| | | | | x | 32 | 32 | x | 3 | 15 |
| | | | | x | (.914) | (.912) | x | (.730) | (.718) |

example, five farms with the highest TE and four farms with the least TE have the same ranks.

**Remark 5.1:** Note that the standard SFM fails to generate the TE scores among farms due to the wrong skewness and see Table 4 and Figure 3 in detail.

### *5.2. Philippines rice production data in 1997*

In this section, we use the Philippines Rice production data in 1997 [25,26]. This dataset does not show wrong skewness and the impact of skew-normal-based SFMs are

**Table 5.** Estimates (with standard error in parentheses) of parameters from $n = 43$ farms in Philippines rice production data in 1997 are given for the standard SFM, SN-SFM and ESN-SFM.

| SFM | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | | | Log-L. | AIC |
|---|---|---|---|---|---|---|---|---|---|---|
| Estimates | −.658 | .387 | .346 | .226 | .173 | .568 | | | −17.215 | 46.43 |
| Standard error | (.838) | (.191) | (.22) | (.145) | (.058) | (.098) | | | | |
| 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 |
| Estimates | −1.113 | .364 | .428 | .198 | .289 | .548 | 3.4 | | −16.885 | 47.769 |
| Standard error | (.937) | (.199) | (.231) | (.139) | (.098) | (.088) | (6.859) | | | |
| ESN-SFM | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\sigma}_v$ | $\hat{\sigma}_u$ | $\hat{\lambda}$ | $\tau$ | Log-L. | AIC |
| Estimates | 1.649 | .516 | .203 | .183 | .971 | .161 | −8.155 | −1.5 | **−15.079** | 46.158 |
| StandardeError | (.751) | (.178) | (.197) | (.315) | (.093) | (.232) | (1.138) | | | |

Note: The log-likelihood (Log-L) and AIC values are also shown.

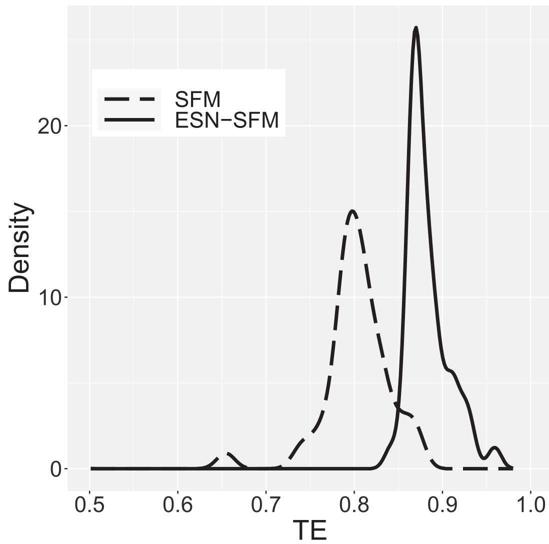


**Figure 4.** Kernel density of efficiency scores for standard SFM and ESN-SFM using Philippines rice production data in 1997.

investigated. Here we estimate the standard SFM, SN-SFM, and ESN-SFM proposed in Section 3. The ML estimates are reported in Table 5.

From Table 5, the AIC value (46.158) in ESN-SFM is lower than those in classical SFM and SN-SFM. This indicates ESN-SFM fits best for these data among the three models. In addition, the skewness parameter estimate $\hat{\lambda} = -8.155$ is significantly less than zero and the left skewness in random error $V$ can be accepted.

Table 6 displays the summary statistics of the efficiency scores of three fitted models. We observe a difference between the ranges of the TE scores in standard SFM (form min $= .471$ to max$= .872$) and ESN-SFM (from min $= .842$ to max$= .959$ ). To better highlight the difference in the distribution of the efficiency scores, the kernel densities of the distribution of efficiency scores in standard SFM and ESN-SFM are given in Figure 4. The standard SFM presents the whole rice industry as less efficient than it is for ESN-SFM. All three models show similar ranks among farms. For instance, from Table 6, No. 32, 31, 38, 19 farms have the highest TE scores and No. 30, 24, 34 farms have the least TE scores under all three models. The difference in the distributions of the TE scores could be due to

**Table 6.** The descriptive statistics of Technical Efficiency for Philippines rice production data in 1997 for the standard SFM, SN-SFM and ESN-SFM are given in the left panel.

| | | | | Highest five farm code | | | Least five farm code | | |
|---|---|---|---|---|---|---|---|---|---|
| | SFM | SN-SFM | ESN-SFM | SFM | SN-SFM | ESN-SFM | SFM | SN-SFM | ESN-SFM |
| Mean | .795 | .709 | .883 | 32 | 32 | 32 | 30 | 30 | 30 |
| S.D. | .063 | .156 | .023 | (.872) | (.903) | (.959) | (.471) | (.105) | (.842) |
| Min | .471 | .105 | .842 | 31 | 31 | 19 | 24 | 24 | 34 |
| Max | .872 | .903 | .959 | (.869) | (.89) | (.933) | (.654) | (.344) | (.86) |
| | | | | 38 | 38 | 38 | 34 | 34 | 24 |
| | | | | (.861) | (.883) | (.927) | (.738) | (.444) | (.86) |
| | | | | 12 | 41 | 31 | 23 | 15 | 15 |
| | | | | (.851) | (.863) | (.924) | (.749) | (.466) | (.861) |
| | | | | 41 | 19 | 2 | 21 | 39 | 39 |
| | | | | (.849) | (.861) | (.914) | (.765) | (.566) | (.863) |

Notes: 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.

the misspecification errors of the underlying economic model. In general, not all variables affecting the efficiency will be included in the data either due to not being well defined or not being measurable. The bias in the economic model, in turn, in TE estimation can flow into the statistical noise, given that the noise term contains all other information that are not included in the model [10]. Thus, a more flexible assumption in noise term in the model, like SN-SFM or ESN-SFM, could blackuce the bias in model and TE estimation.

## 6. Conclusions

In this paper, the classic normal–half-normal SFM is generalized by (extended) skew-normal–half-normal SFM. Technical efficiencies of proposed models are obtained and the skewness of the composite error $\mathcal{E} = V - U$ is derived, showing that both positive and negative skewness are allowed in our new models. Therefore, the proposed models provide an alternative solution to the “wrong skewness problems” of the standard normal–half-normal SFM. For illustration, two real data examples, including the one with the wrong skewness under the classic SFM, show that our (extended) skew-normal-based SFMs perform better than the classic SFM.

In this work, the independent and identically distributed (i.i.d.) random samples are considered. As future work, the proposed models can be extended to the cases given below.

(i) In applied data sets, the i.i.d. assumptions cannot be guaranteed. It is natural to have samples collected to be dependent. We may consider the components of the error term $V$ in SFM are dependent such that their joint distribution is a multivariate skew-normal distribution defined by Azzalini and Dalla Valle [14].

(ii) We may also consider other distributions for the inefficiency term $U$ following gamma, inverse-gamma distributions etc.

**Remark 6.1:** Mathematica [28] codes for calculating skewness in Figure 1 and R [22] codes for rest figures, simulations, and real data analysis are available upon request.

## Acknowledgments

The authors would like to thank Professor Hung T. Nguyen for bringing this interesting topic to our attention and his valuable comments. The authors would also like to thank Associate Editor and reviewers for their valuable suggestions and comments, which have improved the manuscript. The second author would like to thank the support of Elizabeth Gaines Mann Professorship from the University of North Alabama.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

*Xiaonan Zhu* http://orcid.org/0000-0002-9226-8201

## References

[1] Aigner D, Lovell CK, Schmidt P. Formulation and estimation of stochastic frontier production function models. J Econometrics. 1977;6(1):21–37.
[2] Meeusen W, van Den Broeck J. Efficiency estimation from Cobb-Douglas production functions with composed error. Int Econ Rev. 1977;18:435–444.
[3] Coelli T, Rao DP, Battese G. An introduction to efficiency and productivity analysis. Boston (MA): Kluwer Academic Publishers; 1998.
[4] Murillo-Zamorano LR. Economic efficiency and frontier techniques. J Econ Surv. 2004;18(1): 33–77.
[5] Zellner A, Kmenta J, Dreze J. Specification and estimation of Cobb-Douglas production function models. Econometrica J Econom Soc. 1966;34:784–795.
[6] Adcock C, Azzalini A. A selective overview of skew-elliptical and related distributions and of their applications. Symmetry. 2020;12(1):118.
[7] Azzalini A. The skew-normal and related families. Vol. 3. New York: Cambridge University Press; 2013.
[8] Green A, Mayes D. Technical inefficiency in manufacturing industries. Econom J. 1991;101(406):523–538.
[9] Wei Z, Conlon EM, Wang T. Asymmetric dependence in the stochastic frontier model using skew normal copula. Int J Approx Reason. 2021;128:56–68.
[10] Bonanno G, De Giovanni D, Domma F. The 'wrong skewness' problem: a re-specification of stochastic frontiers. J Prod Anal. 2017;47(1):49–64.
[11] Cho CK, Schmidt P. The wrong skew problem in stochastic frontier models when inefficiency depends on environmental variables. Empir Econ. 2020;58(5):2031–2047.
[12] Smith MD. Stochastic frontier models with dependent error components. Econom J. 2008;11(1):172–192.
[13] Azzalini A. A class of distributions which includes the normal ones. Scand J Stat. 1985;46:171–178.
[14] Azzalini A, Dalla Valle A. The multivariate skew-normal distribution. Biometrika. 1996;83(4): 715–726.
[15] Azzalini A, Capitanio A. Statistical applications of the multivariate skew normal distribution. J R Stat Soc Ser B (Stat Methodol). 1999;61(3):579–602.
[16] Wang T, Li B, Gupta AK. Distribution of quadratic forms under skew normal settings. J Multivariate Anal. 2009;100(3):533–545.
[17] González-Farías G, Domínguez-Molina A, Gupta AK. Additive properties of skew normal random vectors. J Statist Plann Inference. 2004;126(2):521–534.
[18] Waldman DM. A stationary point for the stochastic frontier likelihood. J Econom. 1982;18(2):275–279.

[19] Battese GE, Coelli TJ. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. J Econom. 1988;38(3):387–399.
[20] Henningsen A, Toomet O. maxLik: A package for maximum likelihood estimation in R. Comput Stat. 2011;26(3):443–458.
[21] Coelli T, Henningsen A, Henningsen MA. Package 'frontier'; 2020'. (Technical Report).
[22] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available from: http://www.R-project.org/.
[23] Joanes DN, Gill CA. Comparing measures of sample skewness and kurtosis. J R Stat Soc Ser D (Statist). 1998;47(1):183–189.
[24] Akaike H. A new look at the statistical model identification. IEEE Trans Automat Control. 1974;19(6):716–723.
[25] Coelli TJ, Rao DSP, O'Donnell CJ, et al. An introduction to efficiency and productivity analysis. New York: Springer Science & Business Media; 2005.
[26] Henningsen A. Introduction to econometric production analysis with R. Department of Food and Resource Economics, University of Copenhagen; 2014.
[27] Bogetoft P, Otto L. Benchmarking with DEA, SFA, and R. Vol. 157. New York: Springer Science & Business Media; 2010.
[28] Wolfram Research, Inc. Mathematica. Version 11.0. Champaign (IL); 2018.