UTILIZATION OF STUDENT’S T DISTRIBUTION TO HANDLE OUTLIERS IN TECHNICAL EFFICIENCY MEASUREMENT

Rizky Zulkarnain¹, Anik Djuraidah², I Made Sumertajaya², Indahwati²
¹ BPS-Statistics Indonesia
² Department of Statistics, IPB University

e-mail: zulqarnaen@bps.go.id

DOI: 10.14710/medstat.14.1.56-67

Abstract: Stochastic frontier analysis (SFA) is the favorite method for measuring technical efficiency. SFA decomposes the error term into noise and inefficiency components. The noise component is generally assumed to have a normal distribution, while the inefficiency component is assumed to have half normal distribution. However, in the presence of outliers, the normality assumption of noise is not sufficient and can produce implausible technical efficiency scores. This paper aims to explore the use of Student’s t distribution for handling outliers in technical efficiency measurement. The model was applied in paddy rice production in East Java. Output variable was the quantity of production, while the input variables were land, seed, fertilizer, labor and capital. To link the output and inputs, Cobb-Douglas or Translog production functions was chosen using likelihood ratio test, where the parameters were estimated using maximum simulated likelihood. Furthermore, the technical efficiency scores were calculated using Jondrow method. The results showed that Student’s t distribution for noise can reduce the outliers in technical efficiency scores. Student’s t distribution revised the extremely high technical efficiency scores downward and the extremely low technical efficiency scores upward. The performance of model was improved after the outliers were handled, indicated by smaller AIC value.

1. INTRODUCTION

Several methods have been proposed to measure technical efficiency of production units. The methods can be categorized into parametric and nonparametric approaches. Parametric approaches employ specific production function such as Cobb-Douglas or transcendental logarithmic (Translog) production functions, while nonparametric approaches do not use specific production function (Fusco, 2017). Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA) are the examples of nonparametric approaches. Septianto and Widiharih (2010) employed DEA to analyze efficiency of rural bank in Semarang city. While the Deterministic Frontier Analysis (DFA) and Stochastic Frontier Analysis (SFA) are the examples of parametric approaches.

Stochastic frontier model is the frequently preferred method for measuring technical efficiency, especially in agricultural studies (see Heriqbaldi et al., 2015; Mariyono, 2018;
Siagian and Soetjipto, 2020; Sholikah and Kadarmanto, 2020). The primary characteristics of stochastic frontier model is that its error is partitioned into random shock (noise) and inefficiency components (Wheat, 2017). The noise component is generally assumed to have a normal distribution, while the inefficiency component could be exponential, half normal, truncated normal or gamma distributions (Wheat et al., 2019).

The normality assumption of noise is not a problem if the outlier does not exist. If the outliers are present, Wheat et al. (2019) showed that the standard stochastic frontier model can produce exaggerated and implausible spreads of technical efficiency scores. Production frontier in stochastic frontier model consists of deterministic and stochastic parts. The deterministic part is determined by specified production function, while the stochastic part is determined by noise. The deviance of each production units to the production frontier is recorded as technical inefficiency. Therefore, the technical efficiency scores depend on the value of noise as well as the specification of production function. If the noise has outlier problems, the resulted technical efficiency scores could be affected.

This paper aims to utilize Student’s t distribution for handling outlier problems in standard stochastic frontier model. Student’s t distribution is one of the heavy-tailed distribution, so that the chance to cover extreme values is greater than the normal distribution. Wheat et al. (2019) had proposed the use of Student’s t distribution to solve the outlier issues in England highways maintenance costs. However, this approach has not been widely explored in Indonesia, especially in agricultural studies. Zulkarnain and Indahwati (2021) showed that there were outlier issues in rice farming activities and employed Cauchy distribution to handle it. This study uses the same data with Zulkarnain and Indahwati (2021), but with different approaches. This study employs Student’s t distribution for noise component to handle outlier issues. Moreover, this study uses likelihood ratio test to choose the best functional form: Cobb-Douglas or Translog production functions. The inefficiency component is assumed to have half normal distribution, so that the model is called Student’s t – Half Normal model. The performance of Student’s t – Half Normal model is compared to standard stochastic frontier model (Normal – Half Normal model) using Akaike Information Criterion (AIC).

2. LITERATURE REVIEW
2.1. Production Function

Process of production describes the transformation of inputs (capital, labor, energy, material) into output. Production function is the description of this production processes in a mathematical form. In other words, production function represents a quantitative relationship between a bundle of inputs and the maximum possible output (Rasmussen, 2013; Ray and Kumbhakar, 2015). There are two kinds of theoretical production functions that are frequently used in literatures, namely Cobb-Douglas (CD) production function and transcendental logarithmic (Translog) production function. Specification of these production functions are as follows:

Cobb-Douglas (CD) production function:

\[ \ln Y_i = \beta_0 + \sum_{j=1}^{p} \beta_j \ln X_{ij} \]  

Transcendental Logarithmic (Translog) production function:

\[ \ln Y_i = \beta_0 + \sum_{j=1}^{p} \beta_j \ln X_{ij} + \sum_{j=1}^{p} \delta_j \ln (X_{ij})^2 + \sum_{j=1}^{p} \sum_{k>j}^{p} \tau_{jk} (\ln X_{ij})(\ln X_{ik}) \]
where \( Y_i \) is the output of \( i \)th production unit, \( X_{ij} \) is the \( j \)th input of \( i \)th production unit, \( \beta_j \) is the linear component parameter, \( \delta_j \) is the quadratic component parameter, \( \tau_{jk} \) is the interaction component parameter and \( P \) is the number of production inputs.

Translog production function is the extension of Cobb-Douglas production function by incorporating quadratic and interaction components of inputs. Thus, Translog production function is better than Cobb-Douglas production function if the quadratic and interaction components are statistically significant. Conversely, Cobb-Douglas production function is better than Translog production function if the quadratic and interaction components are equal zero. Likelihood ratio test can be used to determine the best production function (Liu et al., 2019):

\[
LR = 2(l_{TR} - l_{CD}) \sim \chi^2_k \tag{3}
\]

where \( LR \) is the log-likelihood ratio statistics that has a chi-square distribution, \( l_{TR} \) is the maximum log-likelihood value of the Translog production function, \( l_{CD} \) is the maximum log-likelihood value of the Cobb-Douglas production function, \( k \) is the number of quadratic and interaction components of inputs. If \( LR > \chi^2_{\alpha,k} \) then Translog production function is better than Cobb-Douglas production function. If \( LR < \chi^2_{\alpha,k} \) then Cobb-Douglas production function is better than Translog production function.

2.2. The Concept of Technical Efficiency

![Figure 1. Technical Efficiency Illustration](image)

Technical efficiency can be measured based on input-oriented or output-oriented basis. Input-oriented measures how much inputs can be proportionally cut down without altering the level of output, while output-oriented measures how much output can be increased for the given inputs (O'Donnell, 2018). Figure 1 illustrates the concept of technical efficiency. Suppose A, B and C are three production units with single output \((y)\) and single input \((x)\), where \( f(x) \) is the production frontier. B and C are fully efficient as they are exactly in production frontier, while A is inefficient. According to input-oriented measure, the technical inefficiency of A is \( CA/OA \) (or technical efficiency of A is \( OC/OA \)). According to output-oriented measure, the technical inefficiency of A is \( AB/IB \) (or technical efficiency of A is \( IA/IB \)).

2.3. Stochastic Frontier Analysis (SFA)

Stochastic frontier model can be expressed as follow (Fusco, 2017):

\[
\ln(y_i) = \ln[f(x_i; \beta)] + v_i - u_i \quad ; i = 1,2, ..., n \tag{4}
\]
where $y_i$ is the output of $i$th production unit, $x_i$ is the inputs vector of $i$th production unit, $\beta$ is the vector of parameters, $\ln[f(x_i; \beta)]$ is the specified production function, $v_i$ is the noise component, $u_i$ is the inefficiency component, $v_i$ and $u_i$ are independent.

If Normal – Half Normal model is used, that is if $v_i$ is normally distributed and $u_i$ has half normal distribution, then the density function of $v_i$ and $u_i$ are as follows:

$$f_v(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{ -\frac{v^2}{2\sigma_v^2} \right\}; -\infty < v < \infty, \sigma_v^2 > 0 \quad (5)$$

$$f_u(u) = \frac{2}{\sqrt{2\pi}\sigma_u} \exp\left\{ -\frac{u^2}{2\sigma_u^2} \right\}; u > 0, \sigma_u^2 > 0 \quad (6)$$

Since $v_i$ and $u_i$ are independent, the joint density of $v_i$ and $u_i$ is as follow:

$$f_{u,v}(u, v) = f_v(v) \cdot f_u(u) = \frac{1}{\pi\sigma_u \sigma_v} \exp\left\{ -\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2} \right\} \quad (7)$$

The composed error $\varepsilon_i$ can be expressed as $\varepsilon_i = v_i - u_i$, thus the equation (6) can be written as:

$$f_{u,\varepsilon}(u, \varepsilon) = \frac{1}{\pi\sigma_u \sigma_v} \exp\left\{ -\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon + u)^2}{2\sigma_v^2} \right\} \quad (8)$$

Marginal density function for $\varepsilon$ can be derived by integrating equation (7) with respect to $u$. Fusco (2017) showed that the marginal density of $\varepsilon$ is as follow:

$$f_\varepsilon(\varepsilon) = \int_0^\infty f_{u,\varepsilon}(u, \varepsilon) \, du = \frac{2}{\sqrt{2\pi}\sigma} \left[ 1 - \Phi\left( \frac{\varepsilon + \lambda}{\sigma} \right) \right] \exp\left\{ -\frac{\varepsilon^2}{2\sigma^2} \right\} \quad (9)$$

where $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \sigma_u / \sigma_v$ and $\Phi(.)$ is the cumulative distribution function of the standard normal distribution. The likelihood ($L$) and log-likelihood functions ($l$) are further constructed as follows:

$$L = \prod_{i=1}^n f_\varepsilon(\varepsilon_i) = \left( \frac{2}{\sqrt{2\pi}\sigma} \right)^n \prod_{i=1}^n \left[ 1 - \Phi\left( \frac{\varepsilon_i + \lambda}{\sigma} \right) \right] \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2 \right\} \quad (9)$$

$$l = \ln(L) = \sum_{i=1}^n \left( \frac{1}{2} \ln\left( \frac{2}{\pi} \right) - \ln(\sigma) + \ln \left[ \Phi\left( -\frac{\lambda\varepsilon_i}{\sigma} \right) \right] - \frac{\varepsilon_i^2}{2\sigma^2} \right) \quad (10)$$

where $L$ is the likelihood function, $l$ is the log-likelihood function, $n$ is the number of data, $\varepsilon_i = \ln(y_i) - \ln[f(x_i; \beta)]$, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \sigma_u / \sigma_v$ and $\Phi(.)$ is the cumulative distribution function of the standard normal distribution.

The parameters $\beta$, $\lambda$ and $\sigma$ are estimated by maximizing likelihood or log-likelihood functions in equation (9) or (10). Solution of equation (9) or (10) can be produced using iteration methods such as Newton-Raphson or Fisher scoring. After the parameter estimates are derived, the technical efficiency scores can be calculated. Jondrow et al. (1982) defined the technical efficiency scores for each production units ($TE_i$) as:

$$TE_i = \exp\left[ -E(u_i | \varepsilon_i) \right] ; i = 1, 2, ..., n \quad (11)$$
According to equation (11), the conditional density function \( u|\varepsilon \) is required to derive technical efficiency scores. If we denote \( \sigma_a = \sigma_u \sigma_v / \sigma, \mu_a = -\varepsilon \sigma_u^2 / \sigma, \lambda = \sigma_u / \sigma_v \) and \( \sigma^2 = \sigma_u^2 + \sigma_v^2 \), then the conditional density function \( u|\varepsilon \) is derived as follows:

\[
f(u|\varepsilon) = \frac{f_{u,\varepsilon}(u, \varepsilon)}{f_\varepsilon(\varepsilon)} = \frac{1}{\pi \sigma_u \sigma_v} \exp \left\{ -\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon + u)^2}{2\sigma_v^2} \right\} / \sqrt{2\pi \sigma}
\]

\[
= \frac{1}{\sqrt{2\pi} (\sigma_u \sigma_v)} \left[ 1 - \Phi \left( \frac{\varepsilon \lambda}{\sigma} \right) \right] \exp \left\{ -\left( \frac{(\sigma_u^2 + \sigma_v^2)u^2}{2\sigma_u^2 \sigma_v^2} \right) - \left( \frac{\varepsilon \sigma_u^2 \sigma_v^2}{2\sigma_v^2} \right) \right\}
\]

\[
= \frac{1}{\sqrt{2\pi} \sigma_u \sigma_v} \left[ 1 - \Phi \left( \frac{\mu_a}{\sigma_v} \right) \right] \exp \left\{ -\frac{(\sigma_u^2 + \sigma_v^2)u^2}{2\sigma_u^2 \sigma_v^2} - \frac{\varepsilon \sigma_u^2 \sigma_v^2}{2\sigma_v^2} \right\}
\]

\[
= \frac{1}{\sqrt{2\pi} \sigma_v} \left[ 1 - \Phi \left( \frac{\mu_a}{\sigma_v} \right) \right] \exp \left\{ -\frac{u^2}{2\sigma_v^2} \right\}
\]

\[
= \frac{\Phi \left( \frac{\mu - \mu_a}{\sigma_v} \right) - \Phi \left( \frac{-\mu_a}{\sigma_v} \right)}{\Phi(\sigma_v) - \Phi(\frac{-\mu_a}{\sigma_v})} ; u > 0
\]

where \( \phi(.) \) and \( \Phi(.) \) denote the probability density function and cumulative distribution function of standard normal distribution, respectively.

Density function \( f(u|\varepsilon) \) in equation (12) is the form of truncated normal distribution with parameters \( \mu = \mu_u, \sigma = \sigma_u, \alpha = -\mu_u / \sigma_u, \) and \( \beta = \infty \). Expected value of a truncated normal random variable is \( \mu + \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \sigma \). Thus, the expected value of \( u|\varepsilon \) is given by:

\[
E(u|\varepsilon) = \mu_u + \frac{\phi(-\mu_u / \sigma_u) - \phi(\infty)}{\Phi(\infty) - \Phi(-\mu_u / \sigma_u)} \sigma_u + \frac{\phi(-\mu_u / \sigma_u) \sigma_u}{1 - \Phi(-\mu_u / \sigma_u)} \sigma_u
\]

\[
= \mu_u + \frac{\phi(-\mu_u / \sigma_u) \sigma_u}{\Phi(\mu_u / \sigma_u)} \sigma_u
\]

Finally, technical efficiency scores in equation (11) can be expressed as:

\[
TE_i = \exp \left\{ -E(u_i|\varepsilon_i) \right\} = \exp \left\{ -\mu_{i\varepsilon} - \sigma_i \frac{\phi(-\mu_{i\varepsilon}/\sigma_i)}{\Phi(\mu_{i\varepsilon}/\sigma_i)} \right\} ; i = 1, 2, ..., n
\]

where \( \mu_{i\varepsilon} = -\varepsilon_i \sigma_u^2 / \sigma^2, \sigma_i = \sigma_u \sigma_v / \sigma, \phi(.) \) is the density function of standard normal distribution and \( \Phi(.) \) is the cumulative distribution function of standard normal distribution.

### 2.4. Student’s t – Half Normal Model

If noise component \( (v_i) \) is assumed to have a non-standardized Student’s t distribution – which includes a scale parameter \( \sigma_v \) – and inefficiency component \( (u_i) \) has half normal distribution (Student’s t – Half Normal model), then the density function of \( v_i \) and \( u_i \) are as follows (Wheat et al., 2019):

---

**Rizky Zulkarnain (Utilization of Student’s t Distribution to Handle Outliers)**
\[ f_v(v) = \frac{\Gamma \left( \frac{a + 1}{2} \right)}{\Gamma \left( \frac{a}{2} \right) \sqrt{\pi a} \sigma_v} \left[ 1 + \frac{1}{a} \left( \frac{v}{\sigma_v} \right)^2 \right]^{\frac{a+1}{2}} e^{-\frac{(v)^2}{2\sigma_v^2}} ; -\infty < v < \infty, \sigma_v > 0, a > 0 \] (15)

\[ f_u(u) = \frac{2}{\sqrt{2\pi} \sigma_u} \exp \left\{ -\frac{u^2}{2\sigma_u^2} \right\} ; u > 0, \sigma_u^2 > 0 \]

where \( a \) is the shape parameter that specifies the degree of kurtosis of the Student’s t distribution, \( \sigma_v \) is the scale parameter and \( \Gamma(.) \) is the gamma function. Since \( v_i \) and \( u_i \) are independent, the joint density of \( v_i \) and \( u_i \) is given by:

\[ f_{u,v}(u,v) = \frac{\Gamma \left( \frac{a + 1}{2} \right)}{\Gamma \left( \frac{a}{2} \right) \sqrt{\pi a} \sigma_v} \left[ 1 + \frac{1}{a} \left( \frac{v}{\sigma_v} \right)^2 \right]^{\frac{a+1}{2}} \frac{2}{\sqrt{2\pi} \sigma_u} \exp \left\{ -\frac{u^2}{2\sigma_u^2} \right\} \] (16)

The composed error \( \varepsilon_i \) can be expressed as \( \varepsilon_i = v_i - u_i \), thus the equation (16) can be written as:

\[ f_{u,e}(u,\varepsilon) = \frac{\Gamma \left( \frac{a + 1}{2} \right)}{\Gamma \left( \frac{a}{2} \right) \sqrt{\pi a} \sigma_v} \left[ 1 + \frac{1}{a} \left( \frac{\varepsilon + u}{\sigma_v} \right)^2 \right]^{\frac{a+1}{2}} \frac{2}{\sqrt{2\pi} \sigma_u} \exp \left\{ -\frac{u^2}{2\sigma_u^2} \right\} \] (17)

Marginal density function for \( \varepsilon \) can be derived by integrating equation (17) with respect to \( u \):

\[ f_{\varepsilon}(\varepsilon) = \int_{0}^{\infty} f_{u,e}(u,\varepsilon) \, du = \mathbb{E}_u[f_{v}(\varepsilon + u)] \] (18)

Equation (18) is not a closed form, so that the solution could not be derived directly. As an alternative, the solution of Equation (18) could be approximated using simulation method (see Train, 2009). Using \( f_v(v) \) and \( f_u(u) \) in Equation (15), the Equation (18) can be expressed as follow:

\[ f_{\varepsilon}(\varepsilon) = \int_{0}^{\infty} f_v(\varepsilon + u) \, f_u(u) \, du = \mathbb{E}_u[f_v(\varepsilon + u)] \] (19)

Thus, \( f_{\varepsilon}(\varepsilon) \) is the expected value of \( f_v(\varepsilon + u) \) with respect to \( u \), where \( u \geq 0 \) is a half normal random variable, \( u \sim N^+ \left( 0, \sigma_u^2 \right) \). If we generate a half normal random variable \( (u_q) \) repeatedly, we can approximate the expected value in equation (19) as an average of simulation data as follow:

\[ \bar{f}_{\varepsilon}(\varepsilon) = \frac{1}{Q} \sum_{q=1}^{Q} f_v(\varepsilon + u_q) ; -\infty < \varepsilon < \infty, u_q > 0 \] (20)

where \( Q \) is the number of simulations and \( u_q \) is generated from a half normal distribution, \( u_q \sim N^+ \left( 0, \sigma_u^2 \right) \). By replacing \( f_v(\varepsilon + u_q) \) with density function of \( f_v(v) \) in equation (15), the equation (20) can be expressed as:
\[
\hat{f}_e(\varepsilon) = \frac{1}{Q} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right) \sqrt{\pi a \sigma_v}} \sum_{q=1}^{Q} \left[ 1 + \frac{1}{a} \left( \varepsilon + u_q \right)^2 \right]^{-\frac{a+1}{2}} \]

Furthermore, the likelihood (\(L\)) and log-likelihood (\(l\)) functions can be derived as:

\[
L = \prod_{i=1}^{n} \hat{f}_e(\varepsilon_i) = \left[ \frac{1}{Q} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right) \sqrt{\pi a \sigma_v}} \right] \prod_{i=1}^{n} \left[ \sum_{q=1}^{Q} \left[ 1 + \frac{1}{a} \left( \varepsilon_i + u_q \right)^2 \right]^{-\frac{a+1}{2}} \right] \]

\[
l = \ln(L) = -n \ln Q + n \ln \left[ \frac{1}{Q} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right) \sqrt{\pi a \sigma_v}} \right] - \frac{n}{2}(\ln \pi + \ln \alpha) - n \ln \sigma_v + \sum_{i=1}^{n} \ln \left[ \sum_{q=1}^{Q} \left[ 1 + \frac{1}{a} \left( \varepsilon_i + u_q \right)^2 \right]^{-\frac{a+1}{2}} \right] \]

where \(n\) is the number of data, \(Q\) is number of simulations, \(a\) is the shape parameter, \(\sigma_v\) is the scale parameter, \(\Gamma(\cdot)\) is the gamma function, \(u_q\) is generated from a half normal distribution \(u_q \sim N^+(0, \sigma_u^2)\), and \(\varepsilon_i = \ln(y_i) - \ln[f(x_i; \beta)]\).

The parameters \(\beta, a, \sigma_u\) and \(\sigma_v\) are estimated by maximizing likelihood or log-likelihood functions in equation (22) or (23). Solution of equation (22) or (23) can be produced using iteration methods such as Newton-Raphson or Fisher scoring.

By following Jondrow et al. (1982) in equation (11), the formula of technical efficiency scores can be derived by firstly constructing \(f(u|\varepsilon)\) as follow:

\[
f(u|\varepsilon) = \frac{f_{u,e}(u, \varepsilon)}{f_e(\varepsilon)} = \frac{f_{u,x}(u, \varepsilon)}{f_e(\varepsilon)} = \frac{f_v(\varepsilon + u) f_u(u)}{f_e(\varepsilon)} ; u > 0 \]

Then, the expected value of \(u|\varepsilon\) is given by:

\[
E(u|\varepsilon) = \int_{0}^{\infty} u f(u|\varepsilon) du = \int_{0}^{\infty} \frac{f_v(\varepsilon + u) f_u(u)}{f_e(\varepsilon)} du
\]

\[
= \frac{1}{f_e(\varepsilon)} E_u[u f_v(\varepsilon + u)] = \frac{\left[ \frac{1}{Q} \sum_{q=1}^{Q} u_q f_v(\varepsilon + u_q) \right]}{f_e(\varepsilon)}
\]

\[
= \frac{\sum_{q=1}^{Q} u_q \left[ 1 + \frac{1}{a} \left( \varepsilon + u_q \right)^2 \right]^{-\frac{a+1}{2}}}{\sum_{q=1}^{Q} \left[ 1 + \frac{1}{a} \left( \varepsilon + u_q \right)^2 \right]^{-\frac{a+1}{2}}}
\]

Finally, technical efficiency scores can be calculated as (Wheat et al., 2019):

\[
TE_i = \exp[-E(u_i|\varepsilon_i)] = \exp \left\{ -\frac{\sum_{q=1}^{Q} u_q \left[ 1 + \frac{1}{a} \left( \varepsilon_i + u_q \right)^2 \right]^{-\frac{a+1}{2}}}{\sum_{q=1}^{Q} \left[ 1 + \frac{1}{a} \left( \varepsilon_i + u_q \right)^2 \right]^{-\frac{a+1}{2}}} \right\}
\]
where $TE_i$ is technical efficiency score of $i$th production unit ($i = 1, 2, ..., n$), $n$ is the number of data, $Q$ is the number of simulations, $a$ is the shape parameter, $\sigma_v$ is the scale parameter, $\varepsilon_i = \ln(y_i) - \ln[f(x_i; \beta)]$ and $u_q$ is generated from a half normal distribution $u_q \sim N^+(0, \sigma_u^2)$.

3. METHODOLOGY

3.1. Data

Data was acquired from the 2017 Cost Structure of Rice Farming Units (SOUT), a survey conducted by Statistics Indonesia (BPS). The data covers about fourteen thousand rice farming units in East Java. Quantity of production is used as output variable, while land, seed, fertilizer, labor and capital are used as input variables. To link the output and input variables, this study uses Cobb-Douglas (CD) and Translog production functions.

Model specification using CD:

$$\ln Production_i = \beta_0 + \beta_1 \ln Land_i + \beta_2 \ln Seed_i + \beta_3 \ln Fertilizer_i + \beta_4 \ln Labour_i + \beta_5 \ln Capital_i + v_i - u_i ; i = 1, 2, ..., n$$  \hspace{1cm} (27)

Model specification using Translog:

$$\ln Production_i = \beta_0 + \beta_1 \ln Land_i + \beta_2 \ln Seed_i + \beta_3 \ln Fertilizer_i + \beta_4 \ln Labour_i + \beta_5 \ln Capital_i + \delta_1 (\ln Land_i)^2 + \delta_2 (\ln Seed_i)^2 + \delta_3 (\ln Fertilizer_i)^2 + \delta_4 (\ln Labour_i)^2 + \delta_5 (\ln Capital_i)^2 + \tau_{12} (\ln Land_i \times \ln Seed_i) + \tau_{13} (\ln Land_i \times \ln Fertilizer_i) + \tau_{14} (\ln Land_i \times \ln Labour_i) + \tau_{15} (\ln Land_i \times \ln Capital_i) + \tau_{23} (\ln Seed_i \times \ln Fertilizer_i) + \tau_{24} (\ln Seed_i \times \ln Labour_i) + \tau_{25} (\ln Seed_i \times \ln Capital_i) + \tau_{34} (\ln Fertilizer_i \times \ln Labour_i) + \tau_{35} (\ln Fertilizer_i \times \ln Capital_i) + \tau_{45} (\ln Labour_i \times \ln Capital_i) + v_i - u_i ; i = 1, 2, ..., n$$  \hspace{1cm} (28)

where $Production_i$ is the production quantity of $i$th production unit (kg), $Land_i$ is the cultivated area of $i$th production unit (m²), $Seed_i$ is the use of seed of $i$th production unit (kg), $Fertilizer_i$ is the use of fertilizer of $i$th production unit (kg), $Labour_i$ is the number of labor of $i$th production unit (person-hours), $Capital_i$ is the capital of $i$th production unit (thousand rupiahs), $v_i$ is the noise component, $u_i$ is the inefficiency component, $v_i$ and $u_i$ are independent.

3.2. Analysis Method

The steps of analysis are:

a. Transform the output and input variables into natural logarithm.

b. Choose the best production function for linking output and inputs. The likelihood ratio test in equation (3) is used as follow:

$H_0$: Cobb-Douglas is better than Translog
$H_1$: Translog is better than Cobb-Douglas

$$LR = 2(l_{TR} - l_{CD})$$

where $LR$ is the log-likelihood ratio statistics, $l_{CD}$ and $l_{TR}$ are the maximum log-likelihood value of the Cobb-Douglas and Translog production functions, respectively. If $LR > \chi^2_{0.05;15}$ then Translog is better than Cobb-Douglas production functions. If $LR < \chi^2_{0.05;15}$ then Cobb-Douglas is better than Translog production functions.

c. Estimate the parameters using Normal – Half Normal and Student’s $t$ – Half Normal models. This paper uses rfrontier STATA package to derive the estimates (Wheat et al., 2019).
d. Calculate technical efficiency scores using equation (14) if Normal – Half Normal model is used and equation (26) if Student’s t – Half Normal model is used.
e. Evaluate the performance of Normal – Half Normal and Student’s t – Half Normal models using Akaike Information Criterion (AIC):

\[ AIC = 2p - 2l \]

where \( p \) is the number of parameters in model and \( l \) is the log-likelihood value of the model.

4. RESULT

Table 1 presents the result of Likelihood Ratio (LR) test for choosing the best production function. Either Normal – Half Normal or Student’s t – Half Normal model produces substantial LR statistics. The LR statistics are larger than \( \chi^2_{0.05; 15} \), thus indicate that Translog production function is better than Cobb-Douglas production function for linking output and input variables.

**Table 1. The Result of Likelihood Ratio (LR) Test**

<table>
<thead>
<tr>
<th></th>
<th>LR statistics</th>
<th>Degree of freedom</th>
<th>( \chi^2_{0.05; 15} )</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal – Half Normal</td>
<td>3430.07</td>
<td>15</td>
<td>25.00</td>
<td>0.00</td>
<td>Reject Ho</td>
</tr>
<tr>
<td>Student’s t – Half Normal</td>
<td>3459.96</td>
<td>15</td>
<td>25.00</td>
<td>0.00</td>
<td>Reject Ho</td>
</tr>
</tbody>
</table>

**Table 2. Estimation Results**

<table>
<thead>
<tr>
<th></th>
<th>Normal – Half Normal</th>
<th>Student’s t – Half Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>Standard error</td>
<td>p-value</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>4.822</td>
<td>0.077</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-0.226</td>
<td>0.013</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-0.026</td>
<td>0.013</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.024</td>
<td>0.013</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>0.015</td>
<td>0.013</td>
</tr>
<tr>
<td>( \beta_5 )</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>( \delta_1 )</td>
<td>0.028</td>
<td>0.004</td>
</tr>
<tr>
<td>( \delta_2 )</td>
<td>-0.031</td>
<td>0.006</td>
</tr>
<tr>
<td>( \delta_3 )</td>
<td>-0.030</td>
<td>0.005</td>
</tr>
<tr>
<td>( \delta_4 )</td>
<td>0.016</td>
<td>0.004</td>
</tr>
<tr>
<td>( \delta_5 )</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>( \tau_{12} )</td>
<td>0.038</td>
<td>0.007</td>
</tr>
<tr>
<td>( \tau_{13} )</td>
<td>0.078</td>
<td>0.007</td>
</tr>
<tr>
<td>( \tau_{14} )</td>
<td>-0.018</td>
<td>0.006</td>
</tr>
<tr>
<td>( \tau_{15} )</td>
<td>0.028</td>
<td>0.005</td>
</tr>
<tr>
<td>( \tau_{23} )</td>
<td>0.000</td>
<td>0.008</td>
</tr>
<tr>
<td>( \tau_{24} )</td>
<td>-0.003</td>
<td>0.007</td>
</tr>
<tr>
<td>( \tau_{25} )</td>
<td>-0.026</td>
<td>0.005</td>
</tr>
<tr>
<td>( \tau_{34} )</td>
<td>-0.002</td>
<td>0.007</td>
</tr>
<tr>
<td>( \tau_{35} )</td>
<td>-0.025</td>
<td>0.005</td>
</tr>
<tr>
<td>( \tau_{45} )</td>
<td>0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>( \sigma_u )</td>
<td>0.244</td>
<td>0.005</td>
</tr>
<tr>
<td>( \sigma_v )</td>
<td>0.730</td>
<td>0.008</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>\infty</td>
<td>-</td>
</tr>
</tbody>
</table>

*Statistically significant at the 5% level
The estimation results of Normal – Half Normal and Student’s t – Half Normal models based on Translog production function are presented in Table 2. Coefficient of $\beta_2$ and $\beta_3$ are insignificant in Normal – Half Normal model but become significant in Student’s t – Half Normal model. Estimate of $\sigma_v$ is smaller than the estimate of $\sigma_u$, either in Normal – Half Normal model or in Student’s t – Half Normal model. It indicates that the deviance of production units to the frontier is more due to inefficiency effect.

Table 3 summarizes the results of technical efficiency score for each production unit and its associated ranking. Normal – Half Normal and Student’s t – Half Normal models produce considerable differences of technical efficiency score and ranking for several production units. For example, the ranking of sixth production unit is 676 in Normal – Half Normal model but becomes 52 in Student’s t – Half Normal model.

### Table 3. Technical Efficiency Score and Technical Efficiency Ranking of Each Production Units Based on Normal – Half Normal and Student’s t – Half Normal models

<table>
<thead>
<tr>
<th>Production Unit</th>
<th>Normal – Half Normal</th>
<th>Student’s t – Half Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technical Efficiency Score</td>
<td>Ranking</td>
</tr>
<tr>
<td>1</td>
<td>0.7952</td>
<td>2197</td>
</tr>
<tr>
<td>2</td>
<td>0.6909</td>
<td>5427</td>
</tr>
<tr>
<td>3</td>
<td>0.6685</td>
<td>6106</td>
</tr>
<tr>
<td>4</td>
<td>0.6570</td>
<td>6410</td>
</tr>
<tr>
<td>5</td>
<td>0.8109</td>
<td>1768</td>
</tr>
<tr>
<td>6</td>
<td>0.8616</td>
<td>676</td>
</tr>
<tr>
<td>7</td>
<td>0.5881</td>
<td>8069</td>
</tr>
<tr>
<td>8</td>
<td>0.6852</td>
<td>5606</td>
</tr>
<tr>
<td>9</td>
<td>0.8179</td>
<td>1575</td>
</tr>
<tr>
<td>10</td>
<td>0.5134</td>
<td>9522</td>
</tr>
<tr>
<td>11</td>
<td>0.7850</td>
<td>2513</td>
</tr>
<tr>
<td>12</td>
<td>0.8469</td>
<td>916</td>
</tr>
<tr>
<td>13</td>
<td>0.8285</td>
<td>1304</td>
</tr>
<tr>
<td>14</td>
<td>0.8386</td>
<td>1077</td>
</tr>
<tr>
<td>15</td>
<td>0.6866</td>
<td>5558</td>
</tr>
<tr>
<td>16</td>
<td>0.6475</td>
<td>6653</td>
</tr>
<tr>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>13993</td>
<td>0.8840</td>
<td>360</td>
</tr>
<tr>
<td>13994</td>
<td>0.5785</td>
<td>8248</td>
</tr>
<tr>
<td>13995</td>
<td>0.6770</td>
<td>5856</td>
</tr>
<tr>
<td>13996</td>
<td>0.7639</td>
<td>3271</td>
</tr>
<tr>
<td>13997</td>
<td>0.6580</td>
<td>6386</td>
</tr>
<tr>
<td>13998</td>
<td>0.7290</td>
<td>4376</td>
</tr>
<tr>
<td>13999</td>
<td>0.4279</td>
<td>10851</td>
</tr>
<tr>
<td>14000</td>
<td>0.7661</td>
<td>3190</td>
</tr>
<tr>
<td>14001</td>
<td>0.7070</td>
<td>4970</td>
</tr>
<tr>
<td>14002</td>
<td>0.9247</td>
<td>12</td>
</tr>
<tr>
<td>14003</td>
<td>0.5347</td>
<td>9142</td>
</tr>
<tr>
<td>14004</td>
<td>0.7433</td>
<td>3928</td>
</tr>
<tr>
<td>14005</td>
<td>0.7325</td>
<td>4267</td>
</tr>
<tr>
<td>14006</td>
<td>0.5654</td>
<td>8510</td>
</tr>
<tr>
<td>14007</td>
<td>0.3393</td>
<td>12421</td>
</tr>
<tr>
<td>14008</td>
<td>0.4210</td>
<td>10985</td>
</tr>
</tbody>
</table>
Figure 2 shows the boxplot of $E(u_i|\varepsilon_i)$, which is used to calculate technical efficiency scores. Normal – Half Normal model produces considerable outliers, which is represented as data points located outside the whiskers (Q1 – 1.5 IQR or Q3 + 1.5 IQR). The outliers are reduced in Student’s t – Half Normal model, thus indicates that Student’s t – Half Normal model can handle the outliers.

The scatterplot of technical efficiency scores from Normal – Half Normal and Student’s t – Half Normal models is depicted in Figure 3. The use of Student’s t distribution for noise component results in a shrinkage of technical efficiency scores at the tails. The extremely low technical efficiency scores in Normal – Half Normal model are revised upward by Student’s t – Half Normal model. Conversely, the extremely high technical efficiency scores in Normal – Half Normal model are revised downward by Student’s t – Half Normal model.

![Figure 2. Boxplot of $E(u_i|\varepsilon_i)$](image1)

![Figure 3. Scatterplot of Technical Efficiency Scores](image2)

The performance of Normal – Half Normal and Student’s t – Half Normal models are compared using Akaike Information Criterion (AIC) in Table 4. The AIC value of Student’s t – Half Normal model is smaller than the Normal – Half Normal model, thus indicates that Student’s t – Half Normal model has better performance than the Normal – Half Normal model. In other words, the performance of stochastic frontier model is improved after the outliers are handled.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal – Half Normal</td>
<td>19406.17</td>
</tr>
<tr>
<td>Student’s t – Half Normal</td>
<td>19312.66</td>
</tr>
</tbody>
</table>

5. CONCLUSION

Utilization of Student’s t distribution in noise component of stochastic frontier model can affect the technical efficiency scores. Student’s t – Half Normal model produces less outliers than the Normal – Half Normal model. The performance of Student’s t – Half Normal model is better than the Normal – Half Normal model if the outliers are present.
REFERENCES


