International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Efficiency Analysis of Panel Data on State-Owned Conventional Commercial Banks Using an Operational Approach with the Stochastic Frontier Analysis Method

Mohammad Fathur Rohman[1], Yuciana Wilandari[2]*, Rita Rahmawati[3]

[1,2,3]Departemen of Statistics, Faculty of Science and Mathematics, Diponegoro University, Semarang, 50275, Indonesia
**yuciana.wilandari@gmail.com*

A B S T R A C T

A state-owned conventional commercial bank is a bank that operates on an interest-based system and has at least 51% of its shares owned by the Republic of Indonesia. Examples include BRI, BNI, Bank Mandiri, Eximbank, and BTN. One way to assess a bank's performance is by measuring its efficiency. There are three common approaches to measuring efficiency: asset-based, intermediation, and operational. The operational approach focuses on managing costs and revenues, using total bank revenue as the output variable, and both interest and non-interest expenses as input variables. Banking efficiency can be analyzed using two main methods: the parametric Stochastic Frontier Analysis (SFA) and the non-parametric Data Envelopment Analysis (DEA). In panel data analysis using SFA, two models—fixed effects and random effects—can be applied, with the Hausman test used to determine the appropriate model. This study analyzes the efficiency of state-owned conventional commercial banks in Indonesia from 2014 to 2023 using the SFA method. The Hausman test indicates that the random effects model is most appropriate. The results show that the banks' efficiency levels were high throughout the study period. Moreover, both input variables—interest and non-interest expenses—significantly influenced total bank revenue.

Keywords: State-Owned Conventional Commercial Bank; Operational Approach; Efficiency Level; Stochastic Frontier Analysis; Fixed Effect; Random Effect

1. Introduction

A conventional commercial bank is a financial institution that operates based on an interest system in conducting business activities related to payment traffic services (OJK, 2019). Law of the Republic of Indonesia No. 19 of 2003 explains that a persero is a state-owned enterprise in the form of a limited liability company, with at least 51% of its shares owned by the state. Based on the definition of a conventional commercial bank and a persero, a state-owned conventional commercial bank can be defined as a financial institution in the form of a bank that applies an interest-based system in payment traffic service activities, established as a limited liability company, and with at least 51% of its shares owned by the Republic of Indonesia.

A conventional commercial bank is a financial institution that operates based on an interest system in conducting business activities related to payment services (OJK, 2019). According to the Law of the Republic of Indonesia No. 19 of 2003, a *persero* is a state-owned enterprise in the form of a limited liability company, with at least 51% of its shares owned by the state. Based on these definitions, a state-owned conventional commercial bank can be defined as a financial institution that applies an interest-based system, operates as a limited liability company, and is majority-owned by the government.

According to the Financial Services Authority (OJK), the banks classified as state-owned conventional commercial banks include Bank Rakyat Indonesia (BRI), Bank Negara Indonesia (BNI), Bank Mandiri, Indonesia Eximbank, and Bank Tabungan Negara (BTN). These institutions are expected to perform strongly, given that the majority of their shares are owned by the state. One key parameter to evaluate a bank's performance is its level of efficiency, as highly efficient banks are expected to minimize the risk of bankruptcy.

Putri and Marwadi (2016) identified three approaches to evaluating bank efficiency: the asset approach, the intermediation approach, and the operational approach. The operational approach focuses on analyzing cost management and income generation to improve performance. In this approach, total bank revenue serves as the output variable, while interest and non-interest expenses are used as input variables.

Two primary methods for estimating efficiency are Stochastic Frontier Analysis (SFA), a parametric approach and Data Envelopment Analysis (DEA), which is non-parametric. Ulkhaq (2021) notes that one of the main advantages of SFA over DEA is its capacity to handle panel data. Originally developed by Aigner et al. (1977), SFA models have been extended by Greene (2005) to accommodate fixed and random effects. The Hausman test is employed to determine the most suitable model. These model effects enable estimation of time-varying inefficiency across firms, from which efficiency scores can be calculated.

Several previous studies have applied SFA in various contexts. Hailu and Tanaka (2015), for instance, assessed the efficiency of Ethiopian manufacturing firms using a random effects SFA with the intermediation approach. Their findings showed that the textile industry had the highest efficiency level. Meanwhile, Mutarindwa et al. (2021) used a fixed effects SFA to analyze cost and profit efficiency in banks across 53 African countries, finding that private banks were more efficient than their state-owned counterparts.

Research on the efficiency of state-owned conventional commercial banks is crucial due to their substantial role in the national economy. This study adopts the operational approach, utilizing the SFA method for panel data analysis. The appropriate model—whether fixed or random effects—is determined using the Hausman test, followed by calculation of efficiency scores. The study examines data from 2014 to 2023 and aims to (1) determine the suitable SFA model, (2) calculate and classify efficiency scores, and (3) identify which input variables significantly influence the output variable.

2. Literature Review

Marsondang et al. (2019) stated that efficiency is a critical indicator in evaluating a company's performance, including banking institutions. A high efficiency level suggests that a company can manage its resources effectively to maximize output and profitability, while a low efficiency level indicates a greater risk of bankruptcy. Wasilah (2018) categorized efficiency levels as shown in Table 1.

Table 1. Efficiency Level Categories

Efisiency Score	Category
$0 - 0{,}39$	Inefficient
$0{,}40 - 0{,}59$	Low
$0{,}60 - 0{,}80$	Moderate
$0{,}81 - 0{,}99$	High
1	Perfect Efficiency

Source: Wasilah (2018)

Using the operational approach, the production function (Cobb-Douglas) is defined as follows (Putri & Marwadi, 2016):

$$\ln(Y_{it}) = \beta_0 + \beta_1 \ln(X1_{it}) + \beta_2 \ln(X2_{it}) + (v_{it} - u_{it})$$

where Y represents total revenue, β represents unknown parameters, $X1$ represents interest expenses, $X2$ represents non-interest expenses, v is the random error beyond the firm's control, and u is the inefficiency term under the firm's control.

Stochastic Frontier Analysis (SFA) is an analytical method used to measure efficiency by comparing the actual output of a production unit (e.g., a bank) to the maximum possible output given its input levels (Sitanggang, 2018). Greene (2005) defines the SFA model as:

$$y_{it}^{*} = \alpha + \boldsymbol{\beta}'\boldsymbol{x}_{it}^{*} + v_{it} - Su_{it};\ i = 1, \dots, N;\ t = 1, \dots, T$$

where $y_{it}^{*} = \ln y_{it}$, with y_{it} representing the output variable for the i-bank at time t; $\boldsymbol{x}_{it}^{*} = \ln \boldsymbol{x}_{it}$, where $\boldsymbol{x}_{it}$ is a $C \times 1$ vector of input variables, with C being the number of input variables; v_{it} denotes the random error term uncontrollable for the i-th bank at time t; u_{it} denotes the random inefficiency term controllable for the i-th bank at time t; S takes the value $+1$ for a production function and -1 for a cost function; $\boldsymbol{\beta}'$ represents the $1 \times C$ slope coefficient vector, where C is the number of input variables; α denotes the intercept coefficient, which is a scalar; $v_{it} \sim N[0, \sigma_v^2]$ means that the v_{it} component is normally distributed with mean zero; and variance σ_v^2, and $u_{it} = |U_{it}|$ where $U_{it} \sim N[0, \sigma_u^2]$ indicates that U_{it} is normally distributed with mean zero and variance σ_u^2. Greene (2005) stated that the calculation of efficiency scores using SFA can be expressed by the following equation:

$$Eff_{it} = \exp(-\hat{u}_{it}) \qquad (1)$$

where Eff_{it} represents the efficiency of the i-th company at time t.

$$\hat{u}_{it} = \left(\frac{\sigma\lambda}{1 + \lambda^2} \times \left(\frac{\phi(a_{it})}{1 - \Phi(a_{it})} - a_{it} \right) \right)$$

$$a_{it} = \frac{S\varepsilon_{it}\lambda}{\sigma}$$

$$\lambda = \frac{\sigma_u}{\sigma_v}$$

Greene (2005) also introduced fixed effect (FE) and random effect (RE) models for panel data in SFA. The FE-SFA assumes firm-specific effects are constant over time, while the RE-SFA assumes these effects are random and uncorrelated with input variables. In the FE-SFA, the model includes a firm-specific intercept.

According to Greene (2005), there are two model effects in SFA that can be used to analyze panel data. These two model effects are the fixed effect SFA and the random effect SFA. The fixed effect SFA is a method for analyzing the efficiency level of a firm by assuming that the individual effects of firms in the panel data are constant over time. On the other hand, the random effect SFA is a method for analyzing firm efficiency by assuming that the individual effects in the panel data are random and uncorrelated with the input variables.

Greene (2005) expressed the fixed effect SFA with the following equation:

$$y_{it}^* = \alpha_i + \boldsymbol{\beta}'\boldsymbol{x}_{it}^* + v_{it} - Su_{it}$$

$$\varepsilon_{it} = y_{it}^* - \alpha_i - \boldsymbol{\beta}'\boldsymbol{x}_{it}^*$$

where α_i represents the cross-sectional specific constant for the i-th bank. Maximum likelihood estimation is used to estimate the parameters of the fixed effect SFA model (α_i and $\boldsymbol{\beta}'$) by maximizing the likelihood function. The fixed effect SFA model is defined by the density function:

$$f(y_{it}^*|x_{i1}^*, x_{i2}^*, \ldots, x_{iT}^*) = \frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right) \qquad (2)$$

Based on the density function in Equation (2), the likelihood function for the fixed effect SFA model is obtained as follows:

$$L = f(y_{i1}^*, y_{i2}^*, \ldots, y_{iT}^*|x_{i1}^*, x_{i2}^*, \ldots, x_{iT}^*) = \prod_{i=1}^{N}\prod_{t=1}^{T}\frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right)$$

The $C \times 1$ parameter vector is rewritten as $\boldsymbol{\gamma} = [\boldsymbol{\beta}, \lambda, \sigma]'$, so that the parameter vector has a dimension of $(C+2) \times 1$. Greene (2005) stated that the gradien ($\boldsymbol{g}$) and Hessian ($\boldsymbol{H}$) of the log-likelihood function for estimating the parameters $\boldsymbol{\gamma}$ and α_i are given by:

$$\boldsymbol{g} = \begin{bmatrix} \frac{\partial \log L}{\partial \beta_1} \\ \vdots \\ \frac{\partial \log L}{\partial \beta_C} \\ \frac{\partial \log L}{\partial \lambda} \\ \frac{\partial \log L}{\partial \sigma} \\ \frac{\partial \log L}{\partial \alpha_1} \\ \vdots \\ \frac{\partial \log L}{\partial \alpha_N} \end{bmatrix} \text{ and } \boldsymbol{H} = \begin{bmatrix} \frac{\partial^2 \log L}{\partial \beta_1 \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \beta_1 \partial \beta_C} & \frac{\partial^2 \log L}{\partial \beta_1 \partial \lambda} & \frac{\partial^2 \log L}{\partial \beta_1 \partial \sigma} & \frac{\partial^2 \log L}{\partial \beta_1 \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \beta_1 \partial \alpha_N} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \log L}{\partial \beta_C \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \beta_C \partial \beta_C} & \frac{\partial^2 \log L}{\partial \beta_C \partial \lambda} & \frac{\partial^2 \log L}{\partial \beta_C \partial \sigma} & \frac{\partial^2 \log L}{\partial \beta_C \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \beta_C \partial \alpha_1} \\ \frac{\partial^2 \log L}{\partial \lambda \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \lambda \partial \beta_C} & \frac{\partial^2 \log L}{\partial \lambda \partial \lambda} & \frac{\partial^2 \log L}{\partial \lambda \partial \sigma} & \frac{\partial^2 \log L}{\partial \lambda \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \lambda \partial \alpha_N} \\ \frac{\partial^2 \log L}{\partial \sigma \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \sigma \partial \beta_C} & \frac{\partial^2 \log L}{\partial \sigma \partial \lambda} & \frac{\partial^2 \log L}{\partial \sigma \partial \sigma} & \frac{\partial^2 \log L}{\partial \sigma \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \sigma \partial \alpha_N} \\ \frac{\partial^2 \log L}{\partial \alpha_1 \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \alpha_1 \partial \beta_C} & \frac{\partial^2 \log L}{\partial \alpha_1 \partial \lambda} & \frac{\partial^2 \log L}{\partial \alpha_1 \partial \sigma} & \frac{\partial^2 \log L}{\partial \alpha_1 \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \alpha_1 \partial \alpha_N} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \log L}{\partial \alpha_N \partial \beta_1} & \cdots & \frac{\partial^2 \log L}{\partial \alpha_N \partial \beta_C} & \frac{\partial^2 \log L}{\partial \alpha_N \partial \lambda} & \frac{\partial^2 \log L}{\partial \alpha_N \partial \sigma} & \frac{\partial^2 \log L}{\partial \alpha_N \partial \alpha_1} & \cdots & \frac{\partial^2 \log L}{\partial \alpha_N \partial \alpha_N} \end{bmatrix}$$

The Newton method with iteration r is used to estimate the parameters in the log-likelihood function.

$$\begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\alpha} \end{pmatrix}_r = \begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\alpha} \end{pmatrix}_{r-1} - \boldsymbol{H}_{r-1}^{-1}\boldsymbol{g}_{r-1}$$

The iteration stops when convergence is achieved, that is, when the norm $\left\| \begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\boldsymbol{\alpha}} \end{pmatrix}_r - \begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\boldsymbol{\alpha}} \end{pmatrix}_{r-1} \right\|$ becomes sufficiently small.

According to Greene (2005), the random effects Stochastic Frontier Analysis (SFA) model is specified by the following equation:

$$y_{it}^* = \alpha + \boldsymbol{\beta}'\boldsymbol{x}_{it}^* + v_{it} - Su_{it} + \omega_i$$

$$\varepsilon_{it} = y_{it}^* - (\alpha + \omega_i) - \boldsymbol{\beta}'\boldsymbol{x}_{it}^*$$

where, ω_i represents the cross-sectional error component specific to bank i. The random effects SFA model is then characterized by the following probability density function:

$$f(y_{it}^*|\omega_i) = \frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right)$$

Simulated Maximum Likelihood (SML) is employed to estimate the parameters of the random effects SFA model. Based on the probability density function of the random effects SFA, the likelihood function is derived as follows:

$$f(y_{1t}^*, \ldots, y_{iT}^*|\omega_i) = \prod_{t=1}^{T}\frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right) \qquad (3)$$

The unconditional likelihood function (L_i) is obtained by integrating out ω_i from Equation (3).

$$L_i = f(y_{1t}^*, \ldots, y_{iT}^*) = \int_{\omega_i}\prod_{t=1}^{T}\frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right)g(\omega_i)d\omega_i$$

The unconditional likelihood function (L_i) is difficult to compute because the integral in the function does not have a closed-form solution. Therefore, the unconditional likelihood function (L_i) is reformulated as follows:

$$L_i = f(y_{1t}^*, \ldots, y_{iT}^*) = E_{\omega_i}\left[\prod_{t=1}^{T}\frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}}{\sigma}\right)\right] \qquad (4)$$

Simulation is required to compute the unconditional likelihood function (L_i) in Equation (4). Greene (2001) stated that the unconditional likelihood function (L_i) in Equation (4) is solved using Monte Carlo Simulation. The unconditional likelihood function (L_i) is simulated M times, resulting in the following function:

$$L_i = \frac{1}{M}\sum_{m=1}^{M}\left[\prod_{t=1}^{T}\frac{2}{\sigma}\emptyset\left(\frac{\varepsilon_{it}|\omega_{im}}{\sigma}\right)\Phi\left(\frac{-S\lambda\varepsilon_{it}|\omega_{im}}{\sigma}\right)\right] \qquad (5)$$

Greene (2001) stated that $M = (NT)^{1+\delta}$ for a positive δ, where δ is a parameter that determines the rate at which the number of simulations M increases as the number of observations (NT) grows. The parameter θ is used to characterize the distribution of ω_i, and ω_{im} denotes the m-th simulated value for firm i. To incorporate θ into Equation (5), ω_i can be expressed as $\omega_i = \theta\omega_{i0}$. If ω_i follows a normal distribution, then θ represents its standard deviation and $\omega_{i0} \sim N[0,1]$. Once the simulated unconditional likelihood function (L_i) is obtained, the log-likelihood function for the random effects SFA model can be expressed as follows:

$\log L(\mathbf{\Theta}) = \sum_{i=1}^{N} \log L_i$, $\mathbf{\Theta} = [\alpha, \beta, \lambda, \sigma, \theta]$

$$\log L(\mathbf{\Theta}) = \sum\nolimits_{i=1}^{N} \log \frac{1}{M} \sum\nolimits_{m=1}^{M} \left[\prod\nolimits_{t=1}^{T} \frac{2}{\sigma} \emptyset\left(\frac{\varepsilon_{it}|\omega_{im}}{\sigma}\right) \Phi\left(\frac{-S\lambda\varepsilon_{it}|\omega_{im}}{\sigma}\right) \right]$$

$\mathbf{\Theta}$ denotes the vector of all parameters to be estimated. To obtain estimates of $\mathbf{\Theta}$, the gradient and the Hessian matrix are employed to locate the maximum point of the likelihood function. The gradien $(\boldsymbol{g})$ and Hessian $(\boldsymbol{H})$ can be expressed as follows:

$$\boldsymbol{g} = \begin{bmatrix} \frac{\partial \log L(\mathbf{\Theta})}{\partial \alpha} \\ \frac{\partial \log L(\mathbf{\Theta})}{\partial \beta_1} \\ \vdots \\ \frac{\partial \log L(\mathbf{\Theta})}{\partial \beta_C} \\ \frac{\partial \log L(\mathbf{\Theta})}{\partial \lambda} \\ \frac{\partial \log L(\mathbf{\Theta})}{\partial \sigma} \\ \frac{\partial \log L(\mathbf{\Theta})}{\partial \theta} \end{bmatrix} \text{ and } \boldsymbol{H} = \begin{bmatrix} \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \alpha \partial \theta} \\ \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_1 \partial \theta} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \beta_C \partial \theta} \\ \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \lambda \partial \theta} \\ \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \sigma \partial \theta} \\ \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \beta_1} & \cdots & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \beta_C} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \lambda} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \sigma} & \frac{\partial^2 \log L(\mathbf{\Theta})}{\partial \theta \partial \theta} \end{bmatrix}$$

The Newton method with iteration r is used to estimate the parameters in the log-likelihood function.

$$(\mathbf{\Theta}')_r = (\mathbf{\Theta}')_{r-1} - \boldsymbol{H}_{r-1}^{-1} \boldsymbol{g}_{r-1}$$

The iteration stops when convergence is achieved, that is, when the norm $\|(\mathbf{\Theta}')_r - (\mathbf{\Theta}')_{r-1}\|$ becomes sufficiently small.

According to Side et al. (2020), the Hausman test aims to choose between the Fixed Effects (FE) and Random Effects (RE) models with the following hypotheses:

H_0 : The model employed is RE

H_1 : The model employed is FE

Test Statistic:

$$\chi^2 = [\widehat{\boldsymbol{\beta}}_{\boldsymbol{FE}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}]' [var(\widehat{\boldsymbol{\beta}}_{\boldsymbol{FE}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}})]^{-1} [\widehat{\boldsymbol{\beta}}_{\boldsymbol{FE}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}]$$

where $\widehat{\boldsymbol{\beta}}_{\boldsymbol{FE}}$ is the vector of slope estimates from the FE model, and $\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}$ is the vector of slope estimates from the RE model. If the test statistic χ^2 exceeds the critical value $\chi^2_{(C,\alpha)}$ or the p-value is less than α, then the null hypothesis H_0 is rejected, indicating that the FE model is the appropriate model to use.

The parameter significance test is applied to determine whether the estimated parameters have a significant impact on the model. There are two types of parameter significance tests: the Joint Test (Wald Test) and the Partial Test. Wicaksono et al. (2023) explain that the Wald Test aims to assess the overall effect of all input variables on the output variable. The hypotheses for the Wald Test can be specified as follows:

H_0 : $\beta_1 = \beta_2 = \cdots \beta_C = 0$

H_1 : At least one $\beta_c \neq 0$ for $c = 1,2,\dots,C$

The Wald Test statistic:

$$w = \widehat{\boldsymbol{\beta}}' \widetilde{\boldsymbol{V}}^{-1} \widehat{\boldsymbol{\beta}}$$

where C represents the number of input variables, $\widehat{\boldsymbol{\beta}}$ denotes the vector of estimated slope coefficients, and $\widetilde{\boldsymbol{V}}^{-1}$ denotes the inverse matrix of the variance-covariance matrix of the slope coefficients. If the test statistic w exceeds the critical value $\chi^2_{(C,\alpha)}$, then the null hypothesis H_0 is rejected, indicating that all input variables have a significant effect on the output variable.

Salsabila et al. (2022) explain that the Partial Test is intended to determine the individual significance of each input variable on the output variable. The hypotheses for the Partial Test are as follows:

H_0 : $\beta_c = 0$ for $c = 1,2,\dots,C$

H_1 : $\beta_c \neq 0$ for $c = 1,2,\dots,C$

The Partial Test equation is as follows:

For the Fixed Effects:

$$t = \frac{\hat{\beta}_c}{se(\hat{\beta}_c)}$$

For the Random Effects:

$$Z = \frac{\hat{\beta}_c}{se(\hat{\beta}_c)}$$

where $\hat{\beta}_c$ represents the estimated coefficient of input variable c, and $se(\hat{\beta}_c)$ denotes the standard error of the estimated coefficient for input variable c. The test criterion is that if $|t| > t_{\left(\frac{\alpha}{2};q-C-1\right)}$ or $|Z| > Z_{\left(\frac{\alpha}{2}\right)}$, then the null hypothesis H_0 is rejected, leading to the conclusion that the individual input variable has a significant effect on the output variable.

3. Research Methodology

This study utilizes secondary data obtained from the annual reports of state-owned conventional commercial banks listed on the Otoritas Jasa Keuangan (OJK) website. The data covers the period from 2014 to 2023. The following steps outline the analysis procedure in this study:

1. Collect data on total revenue, interest expenses, and non-interest expenses of state-owned conventional commercial banks registered with OJK from 2014 to 2023. The collected data are transformed into natural logarithms using Microsoft Excel.
2. Conduct model selection between Fixed Effects (FE) and Random Effects (RE) panel data regression models using the Hausman test (software used: StataMP 17).
3. After determining the appropriate model, proceed with Stochastic Frontier Analysis (SFA) modeling to estimate the parameters of the SFA model. The modeling is performed using StataMP 17.
4. Perform parameter significance tests.
5. Estimate efficiency scores using the selected SFA model and classify the efficiency values. Efficiency calculations are conducted with the assistance of Microsoft Excel.

4. Results And Discussion

The research data consist of observations from 50 state-owned conventional commercial bank, spanning the period from 2014 to 2023. The variables used in this study are interest expenses (x_1), non-interest expenses (x_2), and total revenue (y). Below is the descriptive statistics of the research variables:

Table 2. Descriptive Statistics of Variables (Million Rupiah)

	Total Revenue	Interest Expenses	Non-interest Expenses
Mean	70.729.949,98	18.236.077,58	23.923.513,2
Min	3.898.065	2.082.396	458.543
Max	224.621.779	43.812.507	82.191.967

Before proceeding to the data analysis stage, all variables were transformed into natural logarithmic form in accordance with the production function:

$\ln y = \beta_0 + \beta_1 \ln(x_1) + \beta_2 \ln(x_2) + (v - u)$.

The Hausman test was employed to determine the appropriate model between Fixed Effects (FE) and Random Effects (RE), using the following hypotheses :

H_0 : The model employed is RE

H_1 : The model employed is FE

Significance level $\alpha = 5\%$

Test Statistic, $\chi^2 = 0{,}56$

Critical Region, $\chi^2_{2;0{,}05} = 5{,}99$

The null hypothesis H_0 is rejected if $\chi^2 > \chi^2_{(C,\alpha)}$

Decision and Conclusion: Since $\chi^2 < \chi^2_{(2;0{,}05)}$ ($0{,}56 < 5{,}99$), the null hypothesis H_0 fails to be rejected. Therefore, it can be concluded that the appropriate model for the analysis is the Random Effects (RE) model.

Based on the modeling conducted using StataMP 17, the estimated Random Effects Stochastic Frontier Analysis (SFA) model is as follows:

$\ln \hat{y} = 1{,}848438 + 0{,}542957 \ln(x_1) + 0{,}5422495 \ln(x_2)$

$\sigma_u = 0{,}0038357$

$\sigma_v = 0{,}1337543$

The Wald test was employed to determine whether the input variables jointly have a significant effect on the output variable, using the following hypotheses:

H_0 : All input variables have no significant effect on the output variable

H_1 : All input variables have a significant effect on the output variable

Significance level $\alpha = 5\%$

Test Statistic, $w = 6{,}27 \times 10^{11}$

Critical Region, $\chi^2_{(2;0{,}05)} = 5{,}99$

The null hypothesis H_0 is rejected if $w > \chi^2_{(C,\alpha)}$

Decision and conclusion: Since $w > \chi^2_{(2;0{,}05)}$ ($6{,}27 \times 10^{11} > 5{,}99$), the null hypothesis is rejected. Therefore, it can be concluded that the input variables jointly have a significant effect on the output variable.

The Partial Test was conducted to determine the individual effect or impact of each input variable on the output variable, with the following hypotheses:

H_0 : There is no significant effect of the input variable on the output variable

H_1 : There is a significant effect of the input variable on the output variable

Significance level $\alpha = 5\%$

Test Statistic, $Z_{\ln x_1} = 1447{,}43$ and $Z_{\ln x_2} = 774{,}41$

Critical Region, $Z_{\left(\frac{0{,}05}{2}\right)} = 1{,}96$

The null hypothesis H_0 is rejected if $|Z| > Z_{\left(\frac{\alpha}{2}\right)}$

Decision and Conclusion: Since $\left|Z_{\ln x_1}\right| > Z_{\left(\frac{\alpha}{2}\right)}$ ($1447{,}43 > 1{,}96$) and $\left|Z_{\ln x_2}\right| > Z_{\left(\frac{\alpha}{2}\right)}$ ($774{,}41 > 1{,}96$), the null hypothesis is rejected for both variables. Therefore, it can be concluded that both input variables have a significant partial effect on the output variable.

Based on the equation in Equation (1) and the efficiency classification in Table 1, efficiency scores can be calculated and categorized as presented in Table 3.

Table 3. Score Estimates and Their Categories

Banks	Year	Efisiency Score	Categories
BRI	2014	0,9962346	high
	2015	0,9962294	high
	2016	0,9962453	high
	2017	0,9962512	high
	2018	0,9962454	high
	2019	0,9962374	high
	2020	0,9962180	high
	2021	0,9962719	high
	2022	0,9962998	high
	2023	0,9962618	high
BNI	2014	0,9962501	high
	2015	0,9962529	high
	2016	0,9962422	high
	2017	0,9962310	high

	2018	0,9962299	high
	2019	0,9962152	high
	2020	0,9962260	high
	2021	0,9963017	high
	2022	0,9962963	high
	2023	0,9962447	high
Mandiri	2014	0,9962125	high
	2015	0,9962219	high
	2016	0,9962404	high
	2017	0,9962208	high
	2018	0,9962376	high
	2019	0,9962148	high
	2020	0,9962103	high
	2021	0,9962546	high
	2022	0,9962953	high
	2023	0,9962726	high

Table 4. Score Estimates and Their Categories (Cont.)

Banks	Year	Efisiency Score	Categories
BTN	2014	0,9961725	high
	2015	0,9961904	high
	2016	0,9961847	high
	2017	0,9961834	high
	2018	0,9961772	high
	2019	0,9961559	high
	2020	0,9961673	high
	2021	0,9961737	high
	2022	0,9961957	high
	2023	0,9961867	high
Eximbank	2014	0,9964263	high
	2015	0,9964448	high
	2016	0,9964187	high
	2017	0,9964243	high
	2018	0,9963984	high
	2019	0,9963560	high
	2020	0,9963029	high
	2021	0,9962753	high

2022	0,9962293	high
2023	0,9962464	high

5. Conclusion

Based on the results of the efficiency analysis using panel data from state-owned conventional commercial banks and the application of Stochastic Frontier Analysis (SFA), the following conclusions can be drawn:

1. Based on the Hausman test, the appropriate SFA model is the Random Effects SFA. This model was estimated using StataMP 17, resulting in the following equation:

$$\widehat{ln}y = 1{,}848438 + 0{,}542957 \ln(x_1) + 0{,}5422495 \ln(x_2)$$

2. The estimated efficiency scores for all state-owned conventional commercial banks from 2014 to 2023, using the operational approach with the SFA method, fall within the interval of 0.81 to 0.99. Therefore, it can be concluded that all state-owned conventional commercial banks during this period fall into the high-efficiency category.

Referring to the Wald test results, it can be concluded that interest expenses and non-interest expenses jointly or simultaneously have a significant effect on total revenue. Furthermore, based on the Partial Test, both interest expenses and non-interest expenses also have a significant partial effect on total revenue.

References

Aigner, D., Lovel, C.A.K., Schmidt, P. 1977. Formulation and Estimation of Stochastic Frontier Production Function Model. *Journal of Econometrics 6*, Hal 21 – 37.

Greene, W. 2001. *Fixed and Random Effect in Nonlinear Models*. Department of Economics, Stern School of Business, New York University.

Greene, W. 2005. Fixed and Random Effect in Stochastic Frontier Models. *Journal of Productivity Analysis*, 23, Hal. 7 – 32.

Hailu, K.B., Tanaka, M. 2015. A "True" Random Effect Stochastic Frontier Analysis for Technical Efficiency and Heterogeneity: Evidence from Manufacturing Firms in Ethiopia. *Economic Modelling*, No. 50 : Hal 179 – 192.

Marsondang, A., Purwanto, B., Mulyati, H. 2019. Pengukuran Efisiensi Serta Analisis Faktor Internal dan Eksternal Bank yang Memengaruhinya. *Jurnal Manajemen dan Organisasi (JMO) Vol. 10*, No. 1 : Hal. 48 – 62.

Mutarindwa, S., Siraj, I., Stephan, A. 2021. Ownership and Bank Efficiency in Africa: True Fixed Effect Stochastic Frontier Analysis. *Journal of Financial Stability*, No. 54 : Hal 1 – 12.

[OJK] Otoritas Jasa Keuangan. 2019. *Buku 2 Perbankan Seri Literasi Keuangan Perguruan High*. Jakarta: Tirta Segara.

Putri, A.M., Marwadi, W. 2016. Analisis Efisiensi Bank Persero dan Bank Asing dengan Metode Stochastic Frontier Analysis (SFA) dan Analysis Of Variance (ANOVA) (Studi Kasus pada Bank Persero dan Bank Asing di Indonesia periode tahun 2010 – 2014). *Diponegoro Journal of Management Vol. 5*, No 2 : Hal. 1 – 10.

Salsabila, N.A., Juliarto, H.K., Syawal, A.F., Nohe, D.A. 2019. Analisis Regresi Data Panel pada Ketimpangan Pendapatan Daerah di Provinsi Kalimantan Timur. *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya Terbitan II*, Hal 241 – 253.

Side, S., Sukarna, S., Nurfitrah, R. 2020. Analisis Regresi Panel pada Pemodelan Tingkat Kematian Bayi di Provinsi Sulawesi Selatan. https://ojs.unm.ac.id/JMathCoS/article/view/12574. Diakses : 3 Desember 2024

Sitanggang, Y.F. 2018. *Analisis Efisiensi Teknis Usahatani Cabai Merah Keriting Menggunakan Stochastic Frontier Analysis (SFA) di Desa Mojorejo, Kecamatan Wates, Kabupaten Blitar*. (Skripsi, tidak dipublikasikan). Universitas Brawijaya, Malang.

Ulkhaq, M.M. 2021. Metode Stochastic Frontier Analysis untuk Mengukur Efisiensi di Sektor Pendidikan. *Sainteknol Vol 19*, No. 2 : Hal 65 – 73.

Wasilah, S. 2018. *Analisis Tingkat Efisiensi Bank Umum Syariah (BUS) di Indonesia Periode 2011 – 2015 dengan Menggunakan Non-Parametrik Data Envelopment Analysis (DEA)*. (Skripsi, tidak dipublikasikan). UIN Syarif Hidayatullah, Jakarta.

Wicaksono, M.E., Maruddani, D.A.I., Utami, I.T. 2023. Model Regresi Data Panel Dinamis dengan Estimasi Parameter Arellano-Bond pada Pertumbuhan Ekonomi di Indonesia. *Jurnal Gaussian Vol 12*, No. 2 : Hal 266 – 275.