JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e–ISSN: 2614–8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 21, No. 2, January 2025, pp. 439-455
DOI: 10.20956/j.v21i2.37325

Minimum Variance Quadratic Unbiased Estimation Approach in Panel Data Regression Analysis with Two-Way Error Component Estimation Using Biggers Method

Pendekatan *Minimum Variance Quadratic Unbiased Estimation* dalam Analisis Regresi Data Panel dengan Pendugaan Komponen Galat Dua Arah Menggunakan Metode *Biggers*

Andi Atirah Arumtiwi[1], Raupong Raupong[*2], Siswanto Siswanto[3]

[1,2,3)]*Department of Statistics, Hasanuddin University, Indonesia*
Email: [1]*andiatiraharumtiwi@gmail.com*, [2]*raupong.stat.uh@gmail.com*, [3]*siswanto@unhas.ac.id*
**Corresponding Author*
Received: 17 September 2024, *revised:* 20 November 2024, *accepted:* 29 November 2024

Abstract

Panel data that have missing observations can be known as incomplete panel data. The model used is a two-way error component. The missing data estimation used is the Biggers method. The Biggers method uses a matrix approach and produces more accurate estimates when there is a large amount of missing data. This study aims to form a two-way error component incomplete panel data regression model on Manufacturing Company Stock Return data. The method used for estimating the error variance component is Minimum Variance Quadratic Unbiased Estimation (MIVQUE) with parameter estimation using Maximum Likelihood (ML). The method was applied to IDX data for 10 companies from 2014-2021. The results obtained using the MIVQUE method are $\hat{\sigma}_v^2 = 0.1142$, $\hat{\sigma}_\mu^2 = -0.0107$, and $\hat{\sigma}_\lambda^2 = 0.0068$, for the ML method produces $\hat{\beta}_0 = 0.0304719$ $\hat{\beta}_1 = -0.021107$, and $\hat{\beta}_2 = 0.0087936$. Based on this method, the model obtained is $Y_{ij} = 0.0304719 - 0.021107(X_1)_{ij} + 0.0087936(X_2)_{ij}$, so that if there is an increase in the Debt to Equity Ratio (DER), there is a decrease in the value of stock returns, and vice versa for Net Profit Margin (NPM).

Keywords: Incomplete Panel Data, Two-way Error Component, MIVQUE, ML, Stock Return

Abstrak

Data panel yang memiliki observasi hilang dapat dikenal dengan data panel tidak lengkap. Model yang digunakan merupakan komponen galat dua arah. Pendugaan data hilang yang digunakan adalah metode *Biggers*. Metode Biggers menggunakan pendekatan matriks dan menghasilkan pendugaan yang lebih akurat apabila terdapat data hilang dalam jumlah banyak. Penelitian ini bertujuan untuk membentuk model regresi data panel tidak lengkap komponen galat dua arah pada data *Return* Saham Perusahaan Manufaktur. Metode yang digunakan untuk pendugaan komponen variansi galat adalah *Minimum Variance Quadratic Unbiased Estimation* (MIVQUE) dengan pendugaan parameter menggunakan *Maximum Likelihood* (ML). Metode tersebut diaplikasikan pada data Bursa Efek Indonesia untuk 10 perusahaan periode 2014-2021. Hasil yang didapatkan menggunakan metode MIVQUE yaitu $\hat{\sigma}_v^2 = 0.1142$, $\hat{\sigma}_\mu^2 = -0.0107$, dan $\hat{\sigma}_\lambda^2 = 0.0068$, untuk metode ML menghasilkan $\hat{\beta}_0 = 0.0304719$ $\hat{\beta}_1 = -0.021107$, dan $\hat{\beta}_2 = 0.0087936$. Berdasarkan metode tersebut, model yang didapatkan yaitu $Y_{ij} = 0.0304719 - 0.021107(X_1)_{ij} + 0.0087936(X_2)_{ij}$, sehingga apabila terjadi peningkatan pada *Debt to Equity Ratio* (DER) maka terjadi penurunan nilai *return* saham, dan sebaliknya untuk *Net Profit Margin* (NPM)

Kata kunci: Data Panel Tidak Lengkap, Komponen Galat Dua Arah, MIVQUE, ML, *Return* Saham

1. INTRODUCTION

Regression analysis is a statistical analysis used to evaluate the relationship between one or more variables. The variable is divided into two parts, namely the independent variable and the dependent variable, assuming that the observations are independent of each other. Regression analysis methods have undergone rapid development, this is inseparable from the need for methods in statistics that can help analyze the form of data between units, between times, or combining the two, known as panel data or in the context of regression analysis, referred to as panel data regression models [18].

Panel data is from repeated observations on individual units (objects) and the same cross-section data at different times. Panel data can be used to explain economic problems that are quite complex in terms of time series and cross-section data. In some cases, there is often a problem in data collection, namely the presence of missing observations, meaning that not all individuals are observed in the same time span. The missing data problem is not caused by data unobservable but because the data is not available. If this condition occurs, the panel data can be categorized as incomplete panel data (unbalanced panel) [1].

Incomplete panel data with a two-way error component model refers to data where observations are missing or unavailable and can occur due to various factors. Experiments are often not as expected, and various unexpected constraints may arise, causing the data to be incomplete. Missing data can affect the analysis and cause the treatments not to be orthogonal. Therefore, it is necessary to overcome this by estimating the missing data. Estimating missing data can be done with various methods, such as the Yates, EM Algorithm, and Biggers. The Yates method is commonly used because it produces a more accurate and convergent estimation value compared to other methods [11]. The Biggers method is a refinement of the Yates method with missing data estimation using a matrix approach and higher accuracy in analyzing missing data.

Error component models in incomplete panel data can be estimated using various methods, including Analysis of Variance (ANOVA), Maximum Likelihood (ML), Restricted Maximum Likelihood (REML), and Minimum Variance Quadratic Unbiased Estimation (MIVQUE). ANOVA is a method for decomposing the total variance component in the model and estimating its parameters using Maximum Likelihood (ML). The ML method is a way to estimate regression parameters by maximizing the likelihood function. Meanwhile, the MIVQUE method estimates the variance component, resulting in a minimum, invariant, and unbiased error variance [3]. The error variance obtained through the MIVQUE method can be used to estimate the covariance variance matrix, which is then used as a weight in estimating regression parameters using the ML method.

The two-way error component model is a model in which individual factors and time factors influence errors. Data on these factors can be found in the Indonesia Stock Exchange (IDX) data. In addition, there are often problems in data collection in IDX data. In some conditions, some observations are not available or missing, so they are categorized as incomplete panel data. The approach researchers focus on in this study is the analysis of 10 companies listed on the IDX from 2014 to 2021. at the time of data collection, some companies had not been observed for several years, so there were only 65 observations observed. Based on this data, a panel data regression analysis of the two-way error component model with the MIVQUE method using the Biggers Method estimation is carried out.

2. METHODOLOGY

2.1 Multiple Linear Regression

Multiple regression is an analysis aiming to model and explain the relationship between variables, namely between the dependent and independent variables. Multiple linear regression has more than one independent variable. Multiple linear regression aims to measure the intensity of the relationship between two or more variables. The general form of the multiple linear regression model with parameter β and K independent variables is written in equation (2.1) below:

$$Y_i = \beta_0 + \sum_{k=1}^{K} \beta_j X_{ij} + u_i, i = 1,2,\ldots,N;\ j = 1,2,\ldots,T;\ k = 1,2,\ldots,K \tag{2.1}$$

in matrix notation, written in equation (2.2) as follows:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{u} \tag{2.2}$$

where $\boldsymbol{Y}$ is a vector of dependent variables of size $n \times 1$, $\boldsymbol{\beta}$ is a vector of regression parameters of size $(K+1) \times 1$, $\boldsymbol{u}$ is a vector of errors of size $N \times 1$, and $\boldsymbol{X}$ is a matrix of independent variables of size $N \times (K+1)$.

2.2 Multicollinearity Test

Multicollinearity is a condition of a strong linear relationship between independent variables in a regression model. A multicollinearity test is conducted to identify the correlation between independent variables in multiple regression models. The regression model shows multicollinearity when some or all independent variables are under a linear function with perfect linearity [9].

Indications of multicollinearity can be seen by looking at the Variance Inflation Factor (VIF) and Tolerance values. If the VIF value < 10 and the Tolerance is greater than 0.1, it can be concluded that there is no multicollinearity in the model. The formula for determining the VIF value is in equation (2.3) below [6]:

$$VIF = \frac{1}{tolerance} = \frac{1}{1 - R_j^2} \tag{2.3}$$

where R_j^2 is the coefficient of determination between X_j and the other independent variables.

2.3 Heteroscedasticity Test

Heteroscedasticity is a condition where the variance of the error is not the same for all observations in the regression model, but it is said to be homoscedasticity if the error has the same variance [5]. One of the tests performed for heteroscedasticity is the Glejser test. The Glejser test is a test to determine symptoms of heteroscedasticity in a regression model that occurs by regressing the absolute value of the residual as the dependent variable [5]. The rationale behind this test is that if the significance value of the heteroscedasticity test $\geq 0{,}05$, then it is said that there is homoscedasticity in the data. On the other hand, if the significance value is $< 0{,}05$, it can be concluded that there is heteroscedasticity.

2.4 Autocorrelation Test

Autocorrelation refers to the correlation in the regression model between errors in period t and residuals in the previous period $(t-1)$. The ideal regression model is a model that has no autocorrelation. The autocorrelation test can be done with the Durbin Watson (DW) test [14]. According to [6], the formula for obtaining the DW value is in equation (2.4) below:

$$DW = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2} \tag{2.4}$$

DW test criteria in decision making, namely:

1. If $dU < DW < (4 - dU)$ then there is no autocorrelation.
2. If $DW < dL$ then positive autocorrelation
3. If $DW > (4 - dU)$ then negative autocorrelation
4. If $dL < DW < dU$ or $(4 - dU) < DW < (4 - dL)$ then no decision.

2.5 Analysis of Variance

The Analysis of Variance (ANOVA) method is one of the most widely used methods in estimating variance components. The ANOVA estimator is a moment method type estimator that expresses the sum of the squares of the squares of the errors with their expected values. The resulting equation is solved for the unknown variance component. For a balanced model, the ANOVA estimator is the Best Quadratic Unbiased (BQU) estimator of the variance components [13]. Under the assumption of normal errors, this ANOVA estimator is an unbiased and minimum variance estimator. The ANOVA table can be organized into the total sum of squares, the individual sum of squares, and the sum of squares of errors with degrees of freedom. This method was developed to test the overall significance or significance of the estimated regression and to estimate the additional contribution of the variables explaining the regression.

The ANOVA method is a Best Quadratic Unbiased (BQU) estimation technique for the variance component of a complete panel data model. If the errors are assumed to have a normal distribution, this method can be called an efficient estimator because it is unbiased and has a minimum variance. However, when used for incomplete panel data models, ANOVA can cause biased variance component estimation [7]. The structure of the two-way ANOVA table without interaction is in Table 2.1 below.

Table 2.1 ANOVA Structure

Source of diversity	Degrees of Freedom	Sum of Squares	Mean Squares
Time	$T-1$	SST	$\frac{SST}{T-1}$
Individual	$N-1$	SSI	$\frac{SSI}{N-1}$
Error	$(T-1)\ (N-1)$	SSE	$\frac{SSE}{(T-1)(N-1)}$
Total	$TN-1$	SST	

with the variance component estimator for ANOVA on complete data written in equations (2.5), (2.6), and (2.7) as follows,

$$\sigma_v^2 = MSE \tag{2.5}$$

$$\sigma_\mu^2 = \frac{MSI - MSE}{T} \tag{2.6}$$

$$\sigma_\lambda^2 = \frac{MST - MSE}{N} \tag{2.7}$$

2.6 Metode *Biggers*

The Biggers method is one of the missing data estimators in two-way ANOVA. Missing data estimation is performed on incomplete data to obtain efficient results. Results from observations can often disappear or be deliberately deleted for various reasons. Missing observational data can be caused by unavoidable damage to the experimental unit. Data loss can cause problems in the analysis process because the resulting treatment is no longer balanced in each replicate, so not all treatments can be applied to each replicate/group. The balanced treatment is known as the orthogonal nature of the experimental design. This loss of balance can lead to a loss of competition between treatments in each replicate. Therefore, it is important to estimate the missing data [15].

The Biggers method is a refinement of the Yates method introduced by Bigger (1959) and estimates data with a matrix approach. The Biggers method is used to estimate X_{ij}, i.e. $\hat{X}_{ij}$ The sum of the error squares is determined to be the minimum. The Biggers method equation in matrix form is in equation (2.8) below [10]:

$$\boldsymbol{A}_{k\times k}\boldsymbol{X}_{k\times 1} = \boldsymbol{Q}_{k\times 1} \tag{2.8}$$

where

$\boldsymbol{A}_{k\times k}$: Symmetry matrix of elements $(p-1)(r-1)$ for corresponding times and individuals, $(1-p)$ for corresponding individuals, $(1-r)$ for corresponding times with 1 for others. This matrix is a nonsingular

$\boldsymbol{X}_{k\times 1}$: matrix of missing data

$\boldsymbol{Q}_{k\times 1}$: value matrix $pT_g + rB_h - D$ of the corresponding equation with, T_g total error value of each individual, B_h is the total error value of each time, D is the total number of errors of the whole data [15].

Based on equation (2.8), the following equation (2.9) can be obtained:

$$\hat{\boldsymbol{X}}_{k\times 1} = \boldsymbol{A}^{-1}\boldsymbol{Q} \tag{2.9}$$

The rule that determines the element A is:

a. $(p-1)(r-1)$ for missing observations considered

b. $(1-p)$ for i or individual association
c. $(1-r)$ for j or time association
d. 1 for null association.

2.7 Panel Data Regression

Panel data is a combination of cross-sectional data and time series data. Cross-sectional data is data collected from many individuals at one point in time. Time series data is data collected on an individual over time.

Based on the coverage of panel data, there are two types of panel data, balanced panel data and unbalanced panel data. If each person has the same number of observations over time ($T_1 = T_2 = \ldots = T_N$), the panel data is called balanced panel data (complete panel data), and if the number of observations over time is different for each individual (There is $T_{j\neq}T_{j'}, j \neq j'$), the panel data is called unbalanced panel data (incomplete panel data).

The panel data regression model for complete and incomplete panel data can be expressed in equation (2.10) below:

$$Y_{ij} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{ijk} + u_{ij}\,, i = 1, 2, \ldots, N; j = 1,2, \ldots, T \qquad (2.10)$$

where

Y_{ij} : the value of the dependent variable for the-i and the-t time
β_0 : intercept
X_{ij} : observation of the -k independent variable for the-i and the-t time
β_k : regression coefficient, $k = 1,2, \ldots, K$
u_{ij} : error component in panel data regression model
K : number of independent variables
N : number of individual observations
T : the number of time observations
NT : the number of panel data

The error component u_{ij} in the panel data regression model can be differentiated based on the individual effect and time effect, respectively in equations (2.11) and (2.12) below [2], namely:

1. One-way Error Component Regression Model

$$u_{ij} = \mu_i + v_{ij} \qquad (2.11)$$

2. Two-way Error Component Regression Model

$$u_{ij} = \mu_i + \lambda_j + v_{ij} \qquad (2.12)$$

where

μ_i : Unobserved influence of the-i individual without the influence of time factor
λ_j : Unobserved effect of the-j time without the influence of individual factors
v_{ij} : Totally unknown error of the-i individual and the-j time.

2.8 Two-way Component Incomplete Panel Data Regression Model

In the panel regression model in equation (2.10) and the two-way error component in equation (2.12), an incomplete panel regression model is obtained in equation (2.13) as follows:

$$\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{u}_{ij}, \qquad i = 1, 2, \dots, N_j\,;\ j = 1,2, \dots, T_i \tag{2.13}$$

where $\boldsymbol{Y}_{ij}$ is a vector of observations on the dependent variable for the-i individual at the-j time period; $\boldsymbol{X}_{ij} = (X_{ij1}, X_{ij2}, \dots, X_{ijK})$ is a vector of observations for the-i individual at the-j time on the-k independent variable $(k = 1,2, \dots, K)$; $\boldsymbol{\beta}$ is a vector of size $(K+1) \times 1$ of regression parameters; and $\boldsymbol{u}_{ij}$ is the error component.

Incomplete panel data only observe N_j individuals in period t $\left(2 \leq N_j \leq N\right)$, $j = 1, 2, \dots, T$. The errors in (2.13) are assumed to follow a two-way error component model, which means that differences in individual and time characteristics in the model are accommodated in the errors of the model as in equation (2.12) where each error is assumed to be $\mu_i \sim IIN\left(0, \sigma_\mu^2\right)$, $\lambda_t \sim IIN\left(0, \sigma_\lambda^2\right)$ dan $v_{it} \sim IIN(0, \sigma_v^2)$. Wansbeek & Kapteyn (1989) designed the observation in a way that arranges the order of individuals observed in the first period to the second period sequentially from the beginning to the end. Equation (2.14) can be written in the following vector form:

$$\boldsymbol{u} = \boldsymbol{\Delta}_1\boldsymbol{\mu} + \boldsymbol{\Delta}_2\boldsymbol{\lambda} + \boldsymbol{v} \tag{2.14}$$

with $\boldsymbol{\Delta}_1 = (\boldsymbol{D}_1{}', \boldsymbol{D}_2{}', \dots, \boldsymbol{D}_T{}')'$ and $\boldsymbol{\Delta}_2 = \operatorname{diag}\left(\boldsymbol{D_j}\boldsymbol{\iota}_N\right) = \operatorname{diag}\left(\boldsymbol{\iota}_{\boldsymbol{N_j}}\right)$. The matrix $\boldsymbol{D}_j$ of size $N_j \times N$ is generated from the identity matrix $\boldsymbol{I}_N$ by removing the rows corresponding to unobserved individuals in year j, $\boldsymbol{\iota}_{\boldsymbol{N_j}}$ is a vector of unit elements of size N_j. Based on these assumptions, the variance-covariance matrix $\boldsymbol{\Omega}$ can be written in the form of equation (2.15) below:

$$\begin{aligned} \boldsymbol{\Omega} = E(\boldsymbol{u}\boldsymbol{u}') &= \sigma_\mu^2\boldsymbol{\Delta}_1\boldsymbol{\Delta}_1' + \sigma_\lambda^2\boldsymbol{\Delta}_2\boldsymbol{\Delta}_2' + \sigma_v^2\boldsymbol{I}_n \\ &= \sigma_v^2(\boldsymbol{I}_n + \phi_1\boldsymbol{\Delta}_1\boldsymbol{\Delta}_1' + \phi_2\boldsymbol{\Delta}_2\boldsymbol{\Delta}_2') = \sigma_v^2\boldsymbol{\Sigma}, \\ \phi_1 &= \frac{\sigma_\mu^2}{\sigma_v^2} \text{ dan } \phi_2 = \frac{\sigma_\lambda^2}{\sigma_v^2} \end{aligned} \tag{2.15}$$

with n being the total number of observations $\left(n = \sum_{j=1}^{T} N_j\right)$.

Unlike the complete panel data case, the spectral decomposition of the two-way error component model does not yield a simple transformation Fuller & Battese (1974) in [3]. However, Wansbeek & Kapteyn (1989) derived the inverse of $\boldsymbol{\Omega}$ in equation (2.16) below:

$$\sigma_v^2\boldsymbol{\Omega}^{-1} = \boldsymbol{\Sigma}^{-1} = \boldsymbol{V} - \boldsymbol{V}\boldsymbol{\Delta}_2{\boldsymbol{P}^*}^{-1}\boldsymbol{\Delta}_2'\boldsymbol{V}, \tag{2.16}$$

where

$$\begin{aligned} \boldsymbol{V}_{(n\times n)} &= \boldsymbol{I}_n - \boldsymbol{\Delta}_1{\boldsymbol{\Delta}_N^*}^{-1}\boldsymbol{\Delta}_1' \\ \boldsymbol{P}^*{}_{(T\times T)} &= \boldsymbol{\Delta}_T^* - \boldsymbol{\Delta}_{TN}{\boldsymbol{\Delta}_N^*}^{-1}\boldsymbol{\Delta}_{TN}' \\ \boldsymbol{\Delta}_{N(N\times N)}^* &= \boldsymbol{\Delta}_N + (1/\phi_1)\boldsymbol{I}_N \\ \boldsymbol{\Delta}_{T(T\times T)}^* &= \boldsymbol{\Delta}_T + (1/\phi_2)\boldsymbol{I}_T \end{aligned}$$

with $\boldsymbol{\Delta}_N = \boldsymbol{\Delta}_1'\boldsymbol{\Delta}_1 = \operatorname{diag}(T_i)$, $\boldsymbol{\Delta}_T = \boldsymbol{\Delta}_2'\boldsymbol{\Delta}_2 = \operatorname{diag}\left(N_j\right)$ dan $\boldsymbol{\Delta}_{TN} = \boldsymbol{\Delta}'_2\boldsymbol{\Delta}_1$. T_i is the time the-i individual is observed in the panel with $(2 \leq T_i \leq T)$, $i = 1, 2, \dots, N$ and N_j is the number of individuals for the-j time period observed in the panel with $N_j \leq N$, $j = 1, 2, \dots, T$. The $\boldsymbol{\Omega}^{-1}$ matrix is asymmetric in time and individuals, but this significantly reduces computation time compared to using the numerical inverse of the $\mathbf{\Omega}$ matrix.

2.9 Minimum Variance Quadratic Unbiased Estimation

Minimum Variance Quadratic Unbiased Estimation (MIVQUE) is one of the variance component estimation methods. Rao (1971) proposed a general procedure for variance component estimation that does not require distribution assumptions other than the existence of the first four components. This procedure yields the Minimum Norm Quadratic Unbiased Estimation (MINQUE) of the variance component, under normal error conditions, MINQUE is also the MIVQUE estimation [3]. The estimation of the variance component of the two-way error component using MIVQUE is formulated by equation (2.17) below,

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{S}^{-1}\boldsymbol{u} \tag{2.17}$$

where

$\widehat{\boldsymbol{\theta}}' = (\widehat{\boldsymbol{\sigma}}_v^2, \widehat{\boldsymbol{\sigma}}_\mu^2, \widehat{\boldsymbol{\sigma}}_\lambda^2)$,

$$\boldsymbol{S} = [\boldsymbol{s}_{ij}] = [tr(\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{V}_j\boldsymbol{R})], \qquad i,j = 1,2,3, \tag{2.18}$$

$$\boldsymbol{u} = [\boldsymbol{u}_i] = [\boldsymbol{y}'\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{y}], \qquad i = 1,2,3, \tag{2.19}$$

$$\boldsymbol{R} = \boldsymbol{\Sigma}^{-1}[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}]/\boldsymbol{\sigma}_v^2 \tag{2.20}$$

The matrices $\boldsymbol{S}$, $\boldsymbol{u}$, and $\boldsymbol{R}$ correspond to the matrix $\boldsymbol{V}_i$ i.e. $\boldsymbol{V}_1 = \boldsymbol{I}_n$, $\boldsymbol{V}_2 = \boldsymbol{\Delta}_1\boldsymbol{\Delta}_1'$, $\boldsymbol{V}_3 = \boldsymbol{\Delta}_2\boldsymbol{\Delta}_2'$.

2.10 *Maximum Likelihood*

The method used to find model parameter estimates is the Maximum Likelihood Estimation (MLE) method. The purpose of the maximum likelihood method is to maximize the likelihood function [3]. If it is assumed that the error model is multivariate normally distributed with mean 0 and variance $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is the variance covariance matrix of the response variable $\boldsymbol{Y}$, then the ln-likelihood function can be written as follows,

$$\begin{aligned} L(\boldsymbol{X\beta}, \sigma_v^2\boldsymbol{\Sigma}; \boldsymbol{Y}) &= \prod_{m=1}^{n} f_m(\boldsymbol{Y}) \\ &= (2\pi\sigma_v^2)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma_v^2}(Y - \boldsymbol{X\beta})'^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{Y} - \boldsymbol{X\beta})\right] \end{aligned}$$

with $\boldsymbol{u} = \boldsymbol{Y} - \boldsymbol{X\beta}$, then equation (2.21) is obtained.

$$L(\boldsymbol{X\beta}, \sigma_v^2\boldsymbol{\Sigma}; \boldsymbol{Y}) = \prod_{m=1}^{n} f_m(Y) = (2\pi\sigma_v^2)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}e^{-\frac{1}{2\sigma_v^2}\boldsymbol{u}'\boldsymbol{\Sigma}^{-1}\boldsymbol{u}} \tag{2.21}$$

The explicit solution for the $\boldsymbol{\beta}$ form by maximizing the likelihood function, the parameter estimation equation is obtained which can be written in equation (2.22) [4]. The model parameters are obtained by maximizing the ln-likelihood function by finding the solution of the first derivative of l with respect to $\boldsymbol{\beta}$, namely:

$$\left.\frac{\partial l}{\partial \beta}\right|_{\beta=\widehat{\beta}} = \frac{\partial\left(-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma_v^2) - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2\sigma_v^2}(\boldsymbol{Y} - \boldsymbol{X\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta})\right)}{\partial \beta}$$

$$= \frac{\partial\left(-\frac{n}{2}\ln(2\pi)\right)}{\partial\beta} + \frac{\partial\left(-\frac{n}{2}\ln(\sigma_v^2)\right)}{\partial\beta} + \frac{\partial\left(-\frac{1}{2}\ln|\mathbf{\Sigma}|\right)}{\partial\beta}$$

$$+ \frac{\partial\left(-\frac{1}{2\sigma_v^2}(\boldsymbol{Y}-\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\frac{\partial\left((\boldsymbol{Y}-\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\frac{\partial\left((\boldsymbol{Y}'-(\boldsymbol{X\beta})')\mathbf{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\frac{\partial\left((\boldsymbol{Y}'\mathbf{\Sigma}^{-1}-(\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1})(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\frac{\partial\left(\boldsymbol{Y}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}-\boldsymbol{Y}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}-(\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}\boldsymbol{Y}+(\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\frac{\partial\left(\boldsymbol{Y}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}-2(\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}\boldsymbol{Y}+(\boldsymbol{X\beta})'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}\right)}{\partial\beta}$$

$$= -\frac{1}{2\sigma_v^2}\left(-2\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}+2\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}\right)$$

$$= \frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2} - \frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2}$$

Furthermore, the result of the derivative will be equated to zero as follows,

$$\left.\frac{\partial L}{\partial\beta}\right|_{\beta=\hat{\beta}} = \frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2} - \frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2} = 0$$

$$\frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2} = \frac{\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2}$$

$$\boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{X\beta} = \boldsymbol{X}'\mathbf{\Sigma}^{-1}\boldsymbol{Y}\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{Y}$$

then

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{ML}} = (\boldsymbol{X}'\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{Y} \tag{2.22}$$

where

$$\hat{\mathbf{\Sigma}} = \frac{\hat{\mathbf{\Omega}}}{\sigma_v^2}$$

where

$$\hat{\mathbf{\Omega}} = \hat{\boldsymbol{\sigma}}_v^2\boldsymbol{V}_1 + \hat{\boldsymbol{\sigma}}_\mu^2\boldsymbol{V}_2 + \hat{\boldsymbol{\sigma}}_\lambda^2\boldsymbol{V}_3$$

2.11 Indonesia Stock Exchange

The development of stock prices on the Indonesia Stock Exchange (IDX) is influenced not only by business and economic conditions in Indonesia but also by other countries [12]. The IDX acts as an institution that regulates asset transactions for investment purposes in Indonesia and is responsible for overseeing the operation of the capital market to ensure the capital market functions properly. Manufacturing companies are companies whose shares are actively traded on the floor of the IDX [8]. If the stock return value of manufacturing companies increases, it will attract investors to invest their capital. Several factors affect the stock return value of the company, such as the value of Net Profit Margin (NPM) and the value of Debt to Equity Ratio (DER).

2.12 Methods

The steps taken in analyzing the data are as follows:

1. Form a common effect panel data regression model with the OLS method.
2. Conduct a multicollinearity test by looking at VIF and Tolerance values based on equation (2.3).
3. Heteroscedasticity test using the Glejser test.
4. Autocorrelation test using Durbin Watson test.
5. Determine the value of the error variance component using two-way ANOVA.
   a. Form the errors of multiple regression based on equation (2.2).
   b. Estimating missing data using the Biggers method.
      1) Determining the subscripts of missing data
      2) Determining the elements of matrix A
      3) Calculating the value of matrix Q for each subscript
      4) Estimating missing data based on equation (2.9)
   c. Calculating the value of the two-way error component based on equation (2.5), equation (2.6), and equation (2.7).
6. Form an incomplete panel data regression model based on the panel regression model with two-way error components based on equation (13).
7. Estimating the error variance component using the Minimum Variance Quadratic Unbiased Estimation (MIVQUE) method.
   a. Determine the $\boldsymbol{R}$ matrix using equation (2.20).
   b. Calculate the $\boldsymbol{S}$ matrix using equation (2.18).
   c. Calculate the $\boldsymbol{u}$ matrix using equation (2.19).
   d. Estimating the error variance component using equation (2.17).
8. Estimating incomplete panel data regression parameters using ML method.
   a. Determine the structure of the likelihood function for incomplete panel data regression that follows the multivariate normal distribution.
   b. Determine the ln likelihood function.
   c. Maximizing the ln likelihood function by lowering the function against the parameters.
9. Conclude the results based on the formulation of the problem stated.

3. RESULT AND DISCUSSION

3.1 Estimation of Error Variance Components with the Minimum Variance Quadratic Unbiased Estimator Method

Incomplete panel data regression model with two-way error components is influenced by individual factors and time factors. Incomplete panel data regression linear model with two-way error components in equation (2.14) with the data used. The number of individual companies is 10 and the time is 8 years. The incomplete panel data regression used is a two-way error component model with the error components assumed to be identical, independent, and normally distributed, so that equation (3.1) is written in matrix form according to equations (2.13) and (2.14) as follows:

$$\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\Delta}_1\boldsymbol{\mu}_i + \boldsymbol{\Delta}_2\boldsymbol{\lambda}_j + \boldsymbol{v}_{ij}, \qquad i = 1, 2, \dots, 10\ ;\ j = 1,2, \dots, 8 \tag{3.1}$$

where

$\boldsymbol{\Delta}_1$: matrix of size $n \times N$ and $N_j \leq N$ with N is the number of individuals observed during the observation.

$\boldsymbol{\Delta}_2$: matrix of size $n \times T$ and $T_i \leq T$ with T being the number of observation times observation time

$\boldsymbol{\mu}_i$: $(\mu_1, \mu_2, \dots, \mu_{10})'$

$\boldsymbol{\lambda}_j$: $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_8)'$

$\boldsymbol{v}_{ij}$: $\left(\boldsymbol{v}_{11}, \dots, \boldsymbol{v}_{18}, \dots, \dots, v_{101}, \dots, \boldsymbol{v}_{108_{10}}\right)'$

The variance-covariance matrix $\boldsymbol{\Omega}$ based on equation (2.14) can be written in the form of the following equation,

$$\boldsymbol{\Omega} = E(\boldsymbol{u}\boldsymbol{u}') = \sigma_\mu^2\boldsymbol{\Delta}_1\boldsymbol{\Delta}_1' + \sigma_\lambda^2\boldsymbol{\Delta}_2\boldsymbol{\Delta}_2' + \sigma_v^2\boldsymbol{I}_n$$

The equation is inverted to get the following equation (3.2) which will be used in estimation with the Minimum Variance Quadratic Unbiased Estimator (MIVQUE) method.

$$\sigma_v^2\boldsymbol{\Omega}^{-1} = \boldsymbol{\Sigma}^{-1} = \boldsymbol{V} - \boldsymbol{V}\boldsymbol{\Delta}_2\boldsymbol{P}^{*^{-1}}\boldsymbol{\Delta}_2'\boldsymbol{V}, \tag{3.2}$$

MIVQUE is a method that minimizes the estimation of the error variance component. MIVQUE based on the regression model can be calculated using equation (2.20). Estimation of the error variance component can be done using the following equation.

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{S}^{-1}\boldsymbol{u}$$

3.2 Parameter Estimation with Maximum Likelihood Method

The Maximum Likelihood (ML) method is a method used to estimate parameters in the model. Errors are assumed to be normally distributed with mean 0 and variance covariance σ_v^2. If the errors are identically, independently, and normally distributed $\boldsymbol{\varepsilon} \sim N(0, \sigma_v^2)$ then the dependent variable $\boldsymbol{Y}$ is also identically, independently, and normally distributed $\boldsymbol{Y} \sim N(\boldsymbol{X\beta}, \sigma_v^2)$. After obtaining the likelihood function, then determine the ln likelihood function as follows,

$$\begin{aligned} l &= \ln\left(L(\boldsymbol{X\beta}, \sigma_v^2\boldsymbol{\Sigma}; \boldsymbol{Y})\right) \\ &= \ln\left((2\pi\sigma_v^2)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma_v^2}(Y - \boldsymbol{X\beta})'^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{Y} - \boldsymbol{X\beta})\right]\right) \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma_v^2) - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{(\boldsymbol{Y} - \boldsymbol{X\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta})}{2\sigma_v^2} \end{aligned}$$

The model parameters are obtained by maximizing the ln-likelihood function by finding the solution of the first derivative of l with respect to $\boldsymbol{\beta}$ namely:

$$\left.\frac{\partial l}{\partial \beta}\right|_{\beta=\hat{\beta}} = \frac{\partial\left(-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma_v^2) - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2\sigma_v^2}(\boldsymbol{Y}-\boldsymbol{X\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{\partial \beta}$$

$$= \frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2} - \frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2}$$

Furthermore, the result of the derivative will be equated to zero as follows

$$\left.\frac{\partial L}{\partial \beta}\right|_{\beta=\hat{\beta}} = \frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2} - \frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2} = 0$$

$$\frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X\beta}}{\sigma_v^2} = \frac{\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}}{\sigma_v^2}$$

$$\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X\beta} = \boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Y} \tag{3.3}$$

3.3 Incomplete Panel Data Regression Model with Two-way Error Component on Stock Return Data

The data used is IDX data for 2014-2021 with a total of $n = 65$ observations. The number of companies observed is 10 companies, namely $N = 10$ over a period of 8 years $T = 8$.

For each time period in the data, not all data companies are summarized in the IDX. Therefore, the data includes incomplete panel data and can be formed according to the two-way error component regression model. Based on equations (2.17) and (2.18), the incomplete panel data regression model with two-way error components is obtained as follows:

$$\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\Delta}_1\boldsymbol{\mu}_i + \boldsymbol{u};\ \boldsymbol{u} = \boldsymbol{\Delta}_2\boldsymbol{\lambda}_j + \boldsymbol{v}_{ij}, i = 1, 2, \dots, 10\ ;\ j = 1,2, \dots, 8 \tag{3.4}$$

The panel data model does not have a complete two-way error component vector form and there is a $\boldsymbol{\Delta}$ matrix which is a dummy variable. The matrix consists of $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$ matrices which are $n \times N$ and $n \times T$ respectively.

Based on the data used, the matrix $\boldsymbol{\Delta}_1$ is 65×10 and the matrix $\boldsymbol{\Delta}_2$ is 65×8 so that the matrix $\boldsymbol{\Delta}_1 = (\boldsymbol{D}_1{}', \boldsymbol{D}_2{}', \boldsymbol{D}_3{}', \boldsymbol{D}_4{}', \boldsymbol{D}_5{}', \boldsymbol{D}_6{}', \boldsymbol{D}_7{}', \boldsymbol{D}_8{}')'$ and $\boldsymbol{\Delta}_2 = \text{diag}\left(\boldsymbol{D}_j\iota_{10}\right) = \text{diag}\left(\iota_{N_t}\right)$. The matrix $\boldsymbol{D}_j$ of size $N_j \times 10$ is obtained by removing the rows in the identity matrix I_{10} of size 10×10 corresponding to the-*i* firm that is not observed in the-*j* year.

The matrix $\boldsymbol{\Delta}_1$ is obtained from the arrangement of the matrices $\boldsymbol{D}_{1(6\times10)}, \boldsymbol{D}_{2(8\times10)}, \boldsymbol{D}_{3(9\times10)}, \boldsymbol{D}_{4(10\times10)}, \boldsymbol{D}_{5(9\times10)}, \boldsymbol{D}_{6(8\times10)}, \boldsymbol{D}_{7(7\times10)}, \boldsymbol{D}_{8(8\times10)}$ with the following matrix structure:

$$\Delta_{\mathbf{1}(\mathbf{65}\times10)} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6 \\ D_7 \\ D_8 \end{bmatrix} \tag{3.5}$$

The unit element vectors obtained are $\boldsymbol{\iota}_{6(6\times1)}, \boldsymbol{\iota}_{8(8\times1)}, \boldsymbol{\iota}_{9(9\times1)}, \boldsymbol{\iota}_{10(10\times1)}, \boldsymbol{\iota}_{9(9\times1)}, \boldsymbol{\iota}_{8(8\times1)}, \boldsymbol{\iota}_{7(7\times1)}, \boldsymbol{\iota}_{8(8\times1)}$, so the matrix structure of Δ_2 is as follows:

$$\Delta_{2(65\times8)} = \begin{bmatrix} \iota_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \iota_8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \iota_9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \iota_{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \iota_9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \iota_8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \iota_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \iota_8 \end{bmatrix} \tag{3.6}$$

3.4 Multicollinearity Test

This test is conducted to determine the correlation between independent variables. In this study, the VIF test was used, but before that, calculations were made to obtain tolerance and then calculate the VIF value based on equation (3.7). The calculation to produce the VIF value is as follows.

$$VIF = \frac{1}{0.927} = 1.079 \tag{3.7}$$

Based on the results obtained, the VIF value < 10 means that there is no multicollinearity. So it is known that there is no correlation between the independent variables.

3.5 Heteroscedasticity Test

The heteroscedasticity test is carried out to determine the existence of inequality in the error variance between one observation and another in the regression model. This study used the Glejser test by regressing the absolute value as the dependent variable. Based on the decision criteria, namely, if the significant value is obtained ≥ 0.05, it is said that there is homoscedasticity in the data. After doing the calculation, the results obtained with a significant value of X_1 $0.358 > 0.05$ and X_2 $0.818 > 0.05$ mean that homoscedasticity occurs in the data, so it can be seen that the data has the same error variance.

3.6 Autocorrelation Test

The autocorrelation test determines whether the data correlates with one observation and another. In this study, the Durbin Watson test was used by comparing the DW value obtained from the calculation using equation (2.4) obtained $DW = 2.109$ with the dL and dU values of 1.5355 and 1.6621 respectively. Based on the calculations obtained, it shows that $DW < 4 - dU = 2.109 < 2.3379$, meaning that the data does not have autocorrelation symptoms.

3.7 Estimating Missing Data with the Biggers Method

Incomplete panel data contains missing or unobserved data, so to get good results, missing data estimation is performed. In this study, the missing data estimation used is the Biggers method. The Biggers method uses a matrix approach which is a refinement of the Yates method. Based on the BEI data used, it can be seen that 15 observations are not available or also called missing data. In accordance with equations (2.12) and (2.13) there is a matrix A which has subscript elements of the missing data.

The symmetry matrix elements are $(p-1)(r-1)$ for corresponding time and individual, $(1-p)$ for corresponding individual, $(1-r)$ for corresponding time with 1 for other. This matrix is nonsingular. The value of p is the number of individuals, and r is the number of times. Furthermore, if matrix $\boldsymbol{A}$, has been obtained, then the calculation of the $\boldsymbol{Q}$ matrix value is carried out with the equation $pT_g + rB_h - D$, T_g is the total error value of each individual, B_h is the total error value at each time, D is the total amount of error of the entire data.

If matrix $\boldsymbol{A}$ and matrix $\boldsymbol{Q}$ have been obtained, then the estimation value of missing data can be calculated based on equation (3.8).

$$\widehat{\boldsymbol{X}}_{k\times 1} = \boldsymbol{A}^{-1}\boldsymbol{Q} \tag{3.8}$$

The results of estimating missing data can be seen in Table 3.1.

Table 3. 1 Company Data by Year

Company	Year							
	2014	2015	2016	2017	2018	2019	2020	2021
AISA	0.443	-0.440	0.589	-0.464	0.281	0.647	0.638	-0.315
CINT	-0.124	-0.107	-0.106	0.017	-0.190	-0.007	-0.113	-0.093
DLTA	-0.004	-1.020	-0.071	-0.114	0.167	0.187	-0.392	-0.196
DVLA	0.007	-0.267	0.315	0.083	-0.045	0.112	0.033	0.088
INDF	0.001	-0.255	0.505	-0.064	-0.047	0.126	0.020	0.054
KICI	0.259	-0.069	0.568	0.393	0.627	0.253	0.000	0.310
KINO	-0.202	-0.530	-0.242	-0.335	0.289	0.230	-0.253	-0.302
LMPI	-0.211	-0.381	0.169	0.456	0.114	-0.347	-0.097	-0.046
MBTO	-0.384	-0.338	0.287	-0.300	-0.101	-0.226	-0.054	0.488
PYFA	-0.112	-0.205	0.751	-0.122	0.000	-0.133	-0.002	-0.007

Note: Yellow color indicates the value of the missing data estimate

3.8 Determining the Value of Error Variance Components

The values of the error variance components $\sigma_\mu^2, \sigma_\lambda^2$, dan σ_v^2 are calculated first using the concept of two-way variance analysis. The number of observations used after estimating the value of missing data using the Biggers method is 80 observations. The results of the variance analysis calculation based on table (3.1) are as follows:

Table 3. 2 Results of Variance Analysis Calculation

Source	**Degrees of Freedom**	**Sum of Squares**	**Mean Squares**
Time	7	2.295622	0.327946
Individual	9	1.569317	0.174369
Error	48	3.893283	0.081110
Total	64	7.758222	

Based on the results of table (3.2), the components of error variance are obtained with the following calculations.

$$\sigma_v^2 = 0.0811 \tag{3.9}$$

$$\sigma_\mu^2 = \frac{0.174369 - 0.081110}{8} = 0.0117 \tag{3.10}$$

$$\sigma_\lambda^2 = \frac{0.327946 - 0.081110}{10} = 0.0247 \quad (3.11)$$

3.9 Estimating the Error Variance Component Using the Minimum Variance Quadratic Unbiased Estimation Method

Based on equation (3.2), the two-way error component is incomplete panel data regression. Errors are influenced by individuals and time with error variance components. Information on the error variance component obtained using the concept of variance analysis by estimating missing data using the biggers method will be used to calculate the covariance variance matrix $\boldsymbol{\Omega}$ based on equation (2.15). Furthermore, it will be calculated $\boldsymbol{\Sigma}^{-1}$ using equation (2.16).

The $\boldsymbol{\Sigma}^{-1}$ matrix is a 65×65 matrix and is obtained based on the error variance component information with the two-way variance analysis concept. This value will be used to determine the estimation of error variance components using MIVQUE, which involves matrices $\boldsymbol{R}, \boldsymbol{u}, \boldsymbol{S}$. The results of the matrix $\boldsymbol{R}$ in equation (2.20) is a 65×65 matrix. The results of matrix R that have been calculated using the help of SAS IML software will be used to calculate the u and S matrices contained in equations (2.18) and (2.19) with the following results.

$$\boldsymbol{u} = \begin{bmatrix} 837.98 \\ 109.815 \\ 119.943 \end{bmatrix}$$

$$\boldsymbol{S} = \begin{bmatrix} 7511.19 & 2315.03 & 739.241 \\ 2315.03 & 14401.8 & 51.9539 \\ 739.241 & 51.9539 & 5312.08 \end{bmatrix}$$

the results of the estimation of the error variance component using the MIVQUE method based on equation (2.17) are $\hat{\sigma}_v^2 = 0.1142$, $\hat{\sigma}_\mu^2 = -0.0107$, dan $\hat{\sigma}_\lambda^2 = 0.0068$.

3.10 Estimating Parameters Using the Maximum Likelihood Method

The results of estimating the error variance component using the MIVQUE method then estimate the incomplete panel data regression parameters using the Maximum Likelihood method. The parameter estimation results obtained from equation (3.3) using SAS IML software are as follows,

Table 3.3 Parameter Estimation Results

Parameter	Estimation
$\hat{\beta}_0$	0.0304719
$\hat{\beta}_1$	-0.021107
$\hat{\beta}_2$	0.0087936

Based on table (3.3), the two-way error component incomplete panel data regression model is obtained, namely.

$$Y_{ij} = 0.0304719 - 0.021107(X_1)_{ij} + 0.0087936(X_2)_{ij}$$

The regression model explains that if X_1 and X_2 are constant, there is an increase in stock returns of 0.0304719. If X_1 increases by 1 unit, and X_2 is constant, there will be a decrease of

-0.021107. Meanwhile, if X_2 increases by 1 unit, and X_1 is constant, there is an increase of 0.0087936.

Based on the results obtained, it can be seen that the results of the estimation using the MIVQUE method, one of the components is negative, this can occur due to several factors, different from the results of previous research by Dilla (2017). However, the results of parameter estimation using the ML method produce estimates, one of which is also negative, this also happened in previous research results. Therefore, future research should pay attention and expand the two-way unbalanced error component model to consider regression endogeneity, dynamic specification, sampling imprecision, and serial correlation of disturbances.

4. CONCLUSION

The two-way error component incomplete panel data regression model on the data of stock returns of manufacturing companies on the IDX in 2014-2021 obtained is $Y_{ij} = 0.0304719 - 0.021107(X_1)_{ij} + 0.0087936(X_2)_{ij}$. Based on the regression model, it can be seen that if the DER value (X_1) and the NPM value (X_2) are constant, there will be an increase in stock returns 0.0304719. If DER increases, there is a decrease in the stock return value of 0.021107. Meanwhile, if NPM increases, there is an increase in the stock return value of 0.0087936.

Based on the results obtained, if a company experiences an increase in the company's ability to fulfill its obligations or what is called the *Debt to Equity Ratio* and net profit known as *Net Profit Margin* is constant, there will be a decrease in the level of profit from investment. A company will experience an increase in profits from investment (stock*returns*) if the company's ability to meet all its obligations and its net profit is constant, and if in conditions of increasing net profit and the ability to meet its obligations to pay debt is constant.

REFERENCES

[1] Astuti, R.D.W.I., 2010. Analisis Data Panel Tidak Lengkap Model Komponen Error Dua Arah Dengan Metode Minimum Variance Quadratic Unbiased Estimation (MIVQUE).

[2] Baltagi, B.H., & Chang, Y.J., 1994. Incomplete Panels. A Comparative Study of Alternative Estimators for the Unbalanced One-Way Error Component Regression Model. *Journal of Econometrics*, Vol. 62, No. 2, 67–89.

[3] Baltagi, B.H., Seuck H. S., & Byoung C. J., 2002. A Comparative Study of Alternative Estimators for the Unbalanced One-Way Error Component Regression Model. *Journal of Econometrics*, Vol. 5, No. 2, 480–493.

[4] Dilla, S.U, Andriyana, Y., & Sudartianto, 2017. The Influence of SO4 and NO3 to the Acidity (PH) of Rainwater Using Minimum Variance Quadratic Unbiased Estimation (MIVQUE) and Maximum Likelihood Methods. *AIP Conference Proceedings*, Vol. 1827, No. 3, 1-7.

[5] Fitri, Y.R, Salfadri & Sunreni, 2021. Pengaruh Lingkungan Kerja, Motivasi Dan Disiplin Kerja Terhadap Kinerja Karyawan PT. Semen Padang. *Jurnal Matua*, Vol. 3, No. 1, 197–207.

[6] Gujarati, D. N. 2006. *Basic Econometric*. McGraw-Hill/Irwin, New York.

[7] Jacob, C.A, Sumarja, I.W., & Susilawati, M., 2014. Analisis Model Regresi Data Panel Tidak Lengkap Komponen Galat Dua Arah Dengan Penduga Feasible Generalized Least Square (FGLS). *Jurnal Matematika*, Vol. 4, No. 1, 22–38.

[8] Kencana, D.T., 2021. Pengaruh Manajemen Laba Terhadap Return Saham Dengan Variabel Kontrol Return on Equity Pada Perusahaan Manufaktur Dalam Bursa Efek Indonesia. *TECHNOBIZ : International Journal of Business*, Vol. 4, No. 2, 74-85.

[9] Mardiatmoko, G., 2020. Pentingnya Uji Asumsi Klasik Pada Analisis Regresi Linier Berganda. *Jurnal Ilmu Matematika dan Terapan*, Vol. 14, No. 3, 333–342.

[10] Nasution, M. R., 2019. Penerapan Estimasi Data Hilang Menggunakan Metode Yates Metode Biggers Dan Teknik Sidik Peragam.

[11] Nurhafizah, Gunartha, I.G.E., & Fitriyani, N., 2019. Estimasi Data Hilang Pada Rancangan Split–Block Menggunakan Metode Yates. *Journal of Mathematics: Theory and Applications*, Vol. 1, No. 2, 53–58.

[12] Sari, W.I., 2019. Analisis Pengaruh Inflasi, Suku Bunga SBI, Nilai Tukar Terhadap Return LQ 45 Dan Dampaknya Terhadap Indeks Harga Saham Gabungan (IHSG) Di Bursa Efek Indonesia (BEI). *Jurnal SEKURITAS (Saham, Ekonomi, Keuangan dan Investasi)*, Vol. 3, No. 1, 65-76.

[13] Searle, S.R., 1971. A Biometrics Invited Paper. Topics in Variance Component Estimation. *Biometrics*, Vol. 27, No. 1, 1–76.

[14] Setyarini, A., 2020. Analisis Pengaruh CAR, NPL, NIM, BOPO, LDR Terhadap ROA (Studi Pada Bank Pembangunan Daerah Di Indonesia Periode 2015-2018). *Research Fair Unisri 2019* Vol. 4, No. 1, 282–290.

[15] Widiharih, T., 2007. Estimasi Data Hilang Pada Rancangan Acak Kelompok Lengkap. *Jurnal Matematika UNDIP*, Vol. 10, No. 2, 60–65.

[16] Yendra, R., 2017. Teknik Mengatasi Data Hilang Pada Kasus. *Jurnal Sains Matematika dan Statistika*, Vol. 3, No. 2, 74–80.

[17] Yudiasari, R.S, Sutanto, H.T, Oktaviarina. A., 2014. Penanganan Missing Data Pada Rancangan Blok Random Lengkap. *Jurnal Ilmiah Matematika*, Vol. 3, No. 1, 1-10

[18] Zarkasi, R.N, Surya P., & Sifriyani, 2021. Identifikasi Faktor-Faktor Yang Mempengaruhi Menggunakan Regresi Panel. *Jurnal Ilmu Matematika dan Terapan*, Vol. 15, No. 2, 277–282.