CHAPTER 3
METHODOLOGY

Research methods are scientific approaches used to obtain data for specific purposes and uses (Sugiyono, 2017). Additionally, research methodology can be described as a collection of actions undertaken to investigate research problems, along with the rationale for employing particular procedures or techniques to identify, select, process, and analyze information to understand the problem.

3.1 Research Design

Researchers will detail how the research will be conducted, including the methods used to answer questions and achieve research objectives. According to Bougie and Sekaran (2017), a research design is a plan for collecting, measuring, and analyzing data based on the research questions. This study employs a quantitative method. Sugiyono (2017) explains that quantitative methods involve numerical data and data analysis. This study also investigates causality, known as explanatory research, which aims to determine the extent and nature of cause-and-effect relationships. Figure 3.1 below illustrates the study design, which includes observations, preliminary data collection, problem definitions, research questions, theoretical frameworks, variable measurement, data collection, data analysis, conclusions, and recommendations.

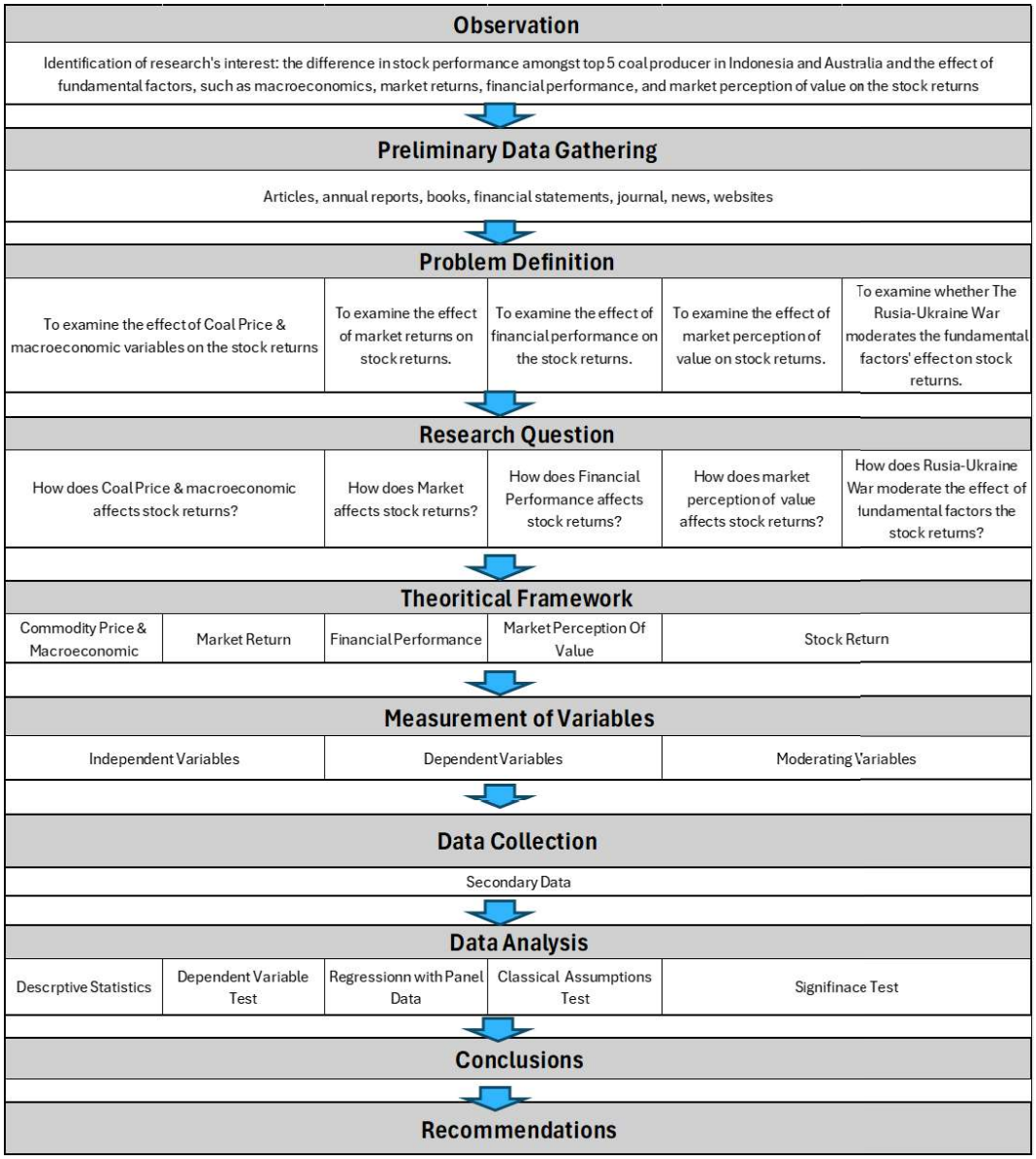

Figure 3.1 Research Design

Source: Author, 2024

3.2 Measurement of Variables

According to Sugiyono (2017), research variables are attributes, characteristics, or values of people, objects, or activities that exhibit certain variations determined by the researcher for study and analysis. This study employs independent, dependent, and moderating variables.

3.2.1 Independent Variable

The independent variable, also known as the causal variable, is responsible for inducing changes in a phenomenon (Kumar, 2011). It can have a positive or negative impact on the dependent variable. The independent variables utilized in this study include:

3.2.1.1 Coal Price Fluctuation Rate

International coal prices are typically measured in US dollars per ton or ton of coal equivalent (tce). Coal export prices are given as 'free on board' (FOB), which includes the cost of coal and transportation from the mine to the export terminal. Import prices are listed as 'cost, insurance, and freight' (CIF), adding the cost of transport to the importer's port to the FOB price. In the USA, the 'free at shipside' (FAS) price is used instead of FOB, excluding loading costs (ECS, 2007). Coal prices vary based on their purpose, characteristics, and market conditions. Historically, coking coal has been more expensive than thermal or steam coal. However, the rising electricity demand in emerging economies has significantly increased the international trade of thermal coal.

In this research coal price will use the index of Global Coal Newcastle Index (GCNI) from 2019-2023. Global Coal, headquartered in London, is an electronic platform that was created in 2001 by coal producers, end-users and others. Trading activities at the Global Coal electronic platform are compiled and published as the Newcastle (NEWC) Price Index, which is based on FOB steam coal prices at the NEWC terminal in Australia. It is an established benchmark for the Asia-Pacific steam coal market.

$$\Delta \mathrm{CPR}t = \frac{\mathrm{CPR}t - \mathrm{CPR}t - 1}{\mathrm{CPR}t - 1} \quad (3.1)$$

Where:

ΔCPR_t=rate of change in the coal price at the period of *t*

CPR_t=coal price at the period of *t*

CPR_{t-1}=coal price the period of $t-1$

3.2.1.2 Exchange Rate

The exchange rate is the price of a country's currency versus another country's currency. The increase or decrease of a country's currency versus another country's currency can be seen through the percentage change in the currency exchange rate. This study will use the quarterly IDR/USD middle rate and AUD/USD issued by Bank Indonesia and issued by the Reserve Bank of Australia (RBA) from 2019 to 2023. The middle rate itself is the average buying rate and selling rate. Exchange rates can be calculated by the formula (Suriyani & Sudiartha, 2018):

$$\Delta \mathrm{ER}_t = \frac{(\mathrm{ER}_t - \mathrm{ER}_{t-1})}{\mathrm{ER}_{t-1}} \quad (3.2)$$

Where:

ΔER_t=rate of change in the exchange rate at the period of *t*

ER_t=exchange rate at the period of *t*

ER_{t-1}=exchange rate at the period of $t-1$

3.2.1.3 Market Return

Market return represents the investors' return for the investments that they have made in the capital market (Thamrin, 2019). This study will use quarterly JKSE returns, IDX Composite or IHSG returns and ASX200 from 2019 to 2023. The formula to calculate market return is:

The formula employed to compute market return is elucidated as follows (Bertuah and Sakti, 2019):

$$MR_t = \frac{IHSG_t - IHSG_{t-1}}{IHSG_{t-1}} \quad (3.3)$$

In the context of the Australian Securities Exchange (ASX), the market return can be computed using a similar formula, adapted to the ASX 200 index. According to research published in the Australian Financial Review (Smith, 2021), the formula employed to compute market return is as follows:

$$MR_t = \frac{ASX200_t - ASX200_{t-1}}{ASX200_{t-1}} \quad (3.3)$$

Where:

MRt = Market return at the period of t

IHSGt = Jakarta Composite Index at the period of t

IHSGt-1 = Jakarta Composite Index at the period of t-1

ASX200t = Australia Composite Index at the period of t

ASX200t-1 = Australia Composite Index at the period of t-1

3.2.1.4 Return on Equity (ROE)

The Return on Equity (ROE) is a metric used to assess a company's profitability, represented as a percentage (%). For instance, a company with a ROA of 10% earns $0.10 for every $1 invested in its assets. ROA evaluates how effectively a company's assets generate income, making a high ROA more favorable to investors than a low one. A higher ROA indicates that asset investments are yielding profitable returns (Malau et al., 2020). This study employs the following measurement scale (Chritianto & Firnanti, 2019), For calculating Return on Equity (ROE), the following formula is used:

$$\mathrm{ROE}_{i,t} = \frac{Net\ Income_{i,t}}{Shareholders\ Equity_{i,t}} \quad (3.4)$$

where,

$ROE_{i,t}$ = ROE of the company i at the period of t Net

$Income_{i,t}$ = Net Income of company i at the period of t

Shareholders' Equityi,t = Shareholders' Equity of company i at the period of t

The result of this formula is also expressed as a percentage. For instance, if a company has an ROE of 15%, it means that for every $1 invested by shareholders, the company generates $0.15 in net income. From an investor's perspective, a higher ROE is preferable to a lower one, as it indicates the company's ability to provide returns to shareholders who have taken on risks with their investments.

3.2.1.5 Current Ratio (CR)

The formula to calculate the Current Ratio (CR) is (Chritianto & Firnanti, 2019):

The liquidity of companies can be measured by using the Current Ratio (CR). According to Kasmir (2019), the current ratio is a measure used to evaluate a company's ability to pay short-term obligations or debts that are due in the near future. In other words, it assesses how much current assets are available to cover short-term liabilities that are soon to mature.

The formula to calculate the Current Ratio (CR) is (Chritianto & Firnanti, 2019):

$$CR_{i,t} = \frac{Current\ Assets_{i,t}}{Current\ Liabilities_{i,t}} \quad (3.5)$$

were,

$CR_{i,t}$ = Current Ratio of the company i at the period of t

$Current\ Assets_{i,t}$ = Current Assets of the company i at the period of t

$Current\ Liabilities_{i,t}$ = Current Liabilities of the company i at the period of t

The current ratio is one of the most widely used liquidity ratios because it provides a straightforward measure of a company’s short-term financial health. A

higher current ratio suggests that the company has more than enough resources to meet its short-term obligations, which is a positive indicator of financial stability. Conversely, a lower current ratio may indicate potential liquidity issues and a higher risk of financial distress. If a company has a CR of 1.5, it means that every $1 of its current liabilities is backed by $1.50 of its current assets. Since the CR reflects how well a company's current assets can cover its current liabilities, a CR of one or higher is considered more favorable by investors compared to a CR below one.

3.2.1.6 Total Asset Turnover (TATO)

The efficiency or activity of companies can be measured by using the Total Asset Turnover (TATO). According to Kurniawan (2021), the formula for total asset turnover is as follows:

$$\text{TATO}_{i,t} = \frac{Sales_{i,t}}{Total\ Assets_{i,t}} \tag{3.6}$$

where,

$\text{TATO}_{i,t}$ = TATO of the company i at the period of t

$\text{Sales}_{i,t}$ = Sales of the company i at the period of t

$\text{Total Assets}_{i,t}$ = Total Assets of the company i at the period of t

If a company has a Total Asset Turnover (TATO) of 1.5, it means that for every $1 invested in assets, the company generates $1.50 in sales. TATO reflects how efficiently a company uses its assets to produce sales. From an investor's perspective, a TATO of one or higher is preferable to a lower TATO. A higher TATO indicates that the company is effectively utilizing its assets to maximize sales, demonstrating strong asset management.

3.2.1.7 Debt To Equity Ratio (DER)

The leverage of companies is often assessed using the Debt-to-Equity Ratio (DER). DER indicates the ratio between a company's debt and its equity. It is measured on a ratio scale with time as the unit. In this research, total debt is also

considered as total liability. The study follows the measurement scale for DER as used by Malau & Murwaningsari (2018) and Chritianto & Firnanti (2019), which is:

$$\text{DER}_{i,t} = \frac{Total\ Debt_{i,t}}{Total\ Equity_{i,t}} \quad (3.7)$$

where,

$DER_{i,t}$ = DER of the company i at the period of t

Total $Debt_{i,t}$ = Total Debt of company i at the period of t

Total $Equity_{i,t}$ = Total Equity of company i at the period of t

If a company has a DER of 0.48, it means that for every $1 of its equity, 48% of that $1 is financed by debt, and 52% is financed by equity. Since the DER indicates the extent to which a company used debt to finance its business operation, the DER that is less than one will be better than the DER that is more than one from investors' perspective.

3.2.1.8 Earning Yield (EY)

The market perception of value of companies can be measured by using the Earning Yield (EY). The formula to calculate the Earning Yield is:

$$EYit = \frac{\text{EPSit}}{\text{Price it}} \quad (3.8)$$

where,

EYi,t = EY of company i at the period of t

EPSit = Earning per share the company i at the period of t

Pricei,t = Price per Share of the company i at the period of t

From a business standpoint, earnings yield improves the precision of performance evaluation over earnings, which may be impacted by earnings management. Executives may be evaluated based on their income, which could lead

them to postpone spending on education and equipment upgrades in order to show bigger incomes. However, as stock prices decline, earnings yield, which also has a negative impact on returns on equity and return on assets, ties earnings to price to reflect earnings inflation. Abraham, Harris, and Auerbach (2017) contend that because earnings are more volatile than dividends and that the change in stock returns is greater than that of dividends, earnings returns should be considered a separate entity.

According to Abraham (2017), earnings yield is the ratio of earnings per share (EPS) to the stock price (E/P). It is the P/E ratio's reciprocal. Thus, expressed as a percentage, earnings yield is equal to EPS / price = 1 / (PER). Investors can immediately determine whether the return is proportionate to the investment risk by looking at the earnings yield. The yield can be used to gauge a stock's rate of return and is a useful Return on Investment (ROI) statistic (David & Randall, 1997).

Table 3. 1 Independent Variable

Variable	Symbol
Coal Price	X_1
Exchange Rate	X_2
Market Return	X_3
Return on Equity (ROE)	X_4
Current Ratio (CR)	X_5
Total Asset Turnover (TATO)	X_6
Debt-to-Equity Ratio (DER)	X_7
Earning Yield (EY)	X_8
The Rusia-Ukraine War	X_9

3.2.2 Dependent Variable

According to Sugiyono (2017), dependent variables are the variables that are often referred to as output variables, criteria, consequent. In Indonesia, a dependent variable is a variable that is affected or that becomes a result because of

independent variables. The dependent variable used in this study is stock performance that consists of stock returns stock risk.

3.2.2.1 Stock Returns

The formula used to measure stock returns, i.e. the current stock price is reduced by the previous period's stock price compared to the last period's stock price. Stock returns is calculated as follows Ristyawan (2019):

$$R_{it} = \frac{P_{it} - P_{it-1}}{P_{it}} \times 100\% \qquad (3.9)$$

where:

Rit = The level of profit of shares i in the period t

Pit = Closing price of shares i in period t (closing/end period)

Pit-1 = Closing price of shares i in the previous period (initial)

3.2.3 Moderating Variable

Moderating variables affect (strengthen and weaken) the relationship between independent and dependent variables (Sugiyono, 2017). The moderation variables in this study were The Rusia-Ukraine War and company value. The formulas used to measure The Rusia-Ukraine War variables are:

Where:

0, t = before 2022 (Jan 2019-Jan 2022) Dt.

1, t = 2022-Dec 2023

3.3 Data Collection

The data sources used in this study are secondary data in the form of macroeconomic data, market return, Coal mining company financial statement data listed on the Indonesia Stock Exchange and Australian Stock Exchange using the quarter report of the period Jan 2019-Dec 2023 taken from the Indonesia Stock Exchange (www.idx.co.id) and Australian Stock Exchange (www.asx.co,au) stock price data from Yahoo Finance data from Jan 2019-Dec 2023

Data collection is carried out by indirect observation by researchers against the object of the study, namely coal mining companies listed on the Indonesia Stock Exchange and Australian Stock Exchange The data collection process is carried out by downloading financial statements report from the Indonesia Stock Exchange website (www.idx.co.id). and Australian Stock Exchange (www.asx.co,au and www.investing.com. Furthermore, researchers recapitulated the data according to the needs of each of the study variables.

3.3.1 Data Collection Procedure

A sample is part of the number and characteristics that a population has. If the population is large and researchers do not study everything (limited funds, energy, and time), researchers can use samples taken from that population (Sugiyono, 2017).

This study used the non-probability sampling technique, which is a sampling technique that does not provide the same opportunity for each element or member of the population to be selected into a sample (Sugiyono, 2017). According to Kumar (2011), non-probability sampling chooses the number of elements in a population that depends on other considerations. There are five commonly used non-probability sampling: quota sampling, accidental sampling, judgmental or purposive sampling, expert sampling, and snowball sampling Kumar (2011).

3.3.2 Research Population and Samples

The term "population" refers to the entire group of people, events, or things of interest in a study. In this research, the population consists of all companies listed on the Indonesia Stock Exchange and Australia Stock Exchange involved in coal mining. The author selected several companies with coal mining as their main business as samples: Top 5 coal producer in Indonesia which are 1. PT Bumi Resources Tbk, (BUMI), 2. PT Adaro Energy Tbk, (ADRO), 3. PT Bayan Resources Tbk, (BYAN), 4.PT Indika Energy Tbk, (INDY), PT Bukit Asam Tbk , (PTBA), and on top 5 coal producer in Australia which are 1. BHP Group Limited

(BHP), 2. Yancoal Limited (YAN), 3. White Haven Coal Limited (WHC), 4. Newhope Coal Limited (NHC), 5. Stanmore Resources Limited (SMR). The criteria for selecting these samples include:

- The company primarily operates in the coal mining sector.
- The company was listed on the Indonesia Stock Exchange before 2019.
- The company has consistently published financial statements from 2019 to 2023
- The company has consistently published share price data from 2019 to 2023

3.4 Techniques of Data Analysis

Data analysis techniques are methods used to convert collected data into valuable information for problem-solving. The process involves four main steps: (1) preparing the data for analysis, (2) understanding the data, (3) testing the data's reliability, and (4) testing the hypotheses formulated (Sekaran & Bougie, 2013). In this study, Microsoft Excel, Eviews, and SPSS Statistics are utilized for data analysis. Eviews provides a range of tools, including statistical and econometric tools, to analyze cross-sectional, time series, and panel data. SPSS Statistics, developed by SPSS Inc., is used for statistical analysis and hypothesis testing to solve research problems. It can also be employed to validate assumptions and derive accurate conclusions.

The data analysis methods applied aim to determine the influence of independent variables on dependent variables. Specifically, the study examines the impact of factors such as Coal Price, exchange rate, market return, return on equity, current ratio, total asset turnover, debt-to-equity ratio, and price-to-earnings ratio on stock returns. Additionally, it investigates the moderating effect of the Rusia-Ukraine War.

3.4.1 Descriptive Analysis

According to Siregar (2013), descriptive analysis is a method of analyzing research data to test generalizations based on a single sample. This approach

involves testing descriptive hypotheses. Sugiyono (2017) explains that descriptive analysis involves examining objects through sample or population data without conducting further analysis or drawing general conclusions.

Haslinda & Jamaluddin (2016) state that descriptive statistics provide a summary or description of the data. This type of statistical analysis is used to gain an overview of the variables in the study. Descriptive statistics involve analyzing the mean values, standard deviations, maximum values, and minimum values to describe the variables. Additionally, they include calculating average frequencies and categorizing statements for descriptive items. Sugiyono (2016) notes that descriptive statistics are employed to analyze data by describing the collected data without aiming to make generalizations. This approach is typically used when analyzing data from entire populations without sampling

3.4.1.1 Mean

Mean is the average of all the data values and a measurement of the data centralization. Mean is also defined as the value obtained by dividing the total values of various given items in a series by the total number of items Kothari (2004). The formula of the mean is:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (3.10)$$

where,

$\overline{X}$ = The symbol for the mean (pronounced as X bar)

X1 + X2 + X3 + Xn = Value of ith item X,

i = 1, 2, …, n

$\sum_{i=1}^{n} X_i$ = = Sum of the value

N = Total number of items

3.4.1.2 Standard Deviation

Standard deviation can be defined as the square root of the average of squares of deviations. Such deviations for the values of individual items in a series

are obtained from the arithmetic average Kothari (2004). The standard deviation is a measure to quantify the amount of variation or dispersion of the data set. The formula of standard deviation is:

$$\sigma = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}} \tag{3.11}$$

where,

σ = The symbol for standard deviation (pronouncedas sigma)

$\bar{X}$ = Mean

Xi = value of i^{th} item

X n = Total number of items

3.4.1.3 Correlation Coefficient

The correlation coefficient measures the strength of the linear relationship between the relative movements of two variables (Ganti, 2020). The formula of correlation is:

$$r_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y} \tag{3.12}$$

where,

rxy = correlation of variables x and y

Cov(x,y) = covariance between x and y

σx = standard deviation of x

σy = standard deviation of y

Value of the correlation coefficient range between -1.0 and 1.0; when the correlation value is -1.0, it means the correlation is a perfect negative correlation. A perfect negative correlation indicates the two variables will move in the opposite direction, in which as one variable increases, the other variable will decrease. Meanwhile, when the correlation value is 1.0, it means the correlation is a perfect positive correlation. A perfect positive correlation indicates the two variables will

move in the same direction, in which as one variable increases, the other variable will also increase. When the value of the correlation is 0.0, there is no linear relationship between the movement of thetwo variables. In addition, when the value of the correlation is greater than 1.0 orless than -1.0, there was an error in the correlation measurement.

The degree of the strength of the relationship between the two variables is based on the correlation coefficient value. For instance, if a correlation value is 0.4, the correlation is a positive correlation between two variables; however, the correlation is weak. The correlation coefficient can be considered a strong correlation if the value of the correlation coefficient is greater than 0.8 or less than -0.8 (Ganti, 2020). The correlation coefficient can be shown in a table, which is called a correlation matrix.

The correlation coefficient uses the following hypothesis to perform testing:

$$H_0: \ r_{ij} = 0$$

$$Ha: \ rij \neq 0$$

Two criteria are used to measure the hypotheses mentioned above:

a. If p-value < significance level of 0.05, the H_0 is rejected = There is asignificant linear relationship or correlation between I and j
b. If p-value > significance level of 0.05, the H_0 is is not rejected = There is no significant linear relationship or correlation between I and j

3.4.2 Panel Data Regression Analysis

Analysis of panel data in research is done by regression of panel data. The regression equations in this study are as follows:

$$Y_{it} = \alpha + \beta 1X_{1t} - \beta 2X_{2t} + \beta 3X_{3t} + \beta 4X_{4it} + \beta 5X_{5it} + \beta 6X_{6it} - \beta 7X_{7it} - \beta 8X_{8it} + \beta 9D_t + \beta 10X_{1t}*D_t + \beta 11X_{2t}*D_t + \beta 12X_{3t}*D_t + \beta 13X_{4it}*D_t + \beta 14X_{5it}*D_t + \beta 15X_{6it}*D_t + \beta 16X_{7it}*D_t + \beta 17X_{8it}*D_t + \varepsilon_{it} \quad (3.13)$$

where :

Y_{it} = Stock Returns of the company i at period t

X_{1t} = Coal Price at the period of t

X_{2t} = Changes in the currency exchange rate of IDR or AUD versus USD at the period oft

X_{3t} = The return of IDX or ASX at the period of t

X_{4t} = Return on Equity (ROE) of the company i at period t

X_{5it} = Current Ratio (CR) of the company i at period t

X_{6it} = Total Asset Turnover (TATO) of the company i at period t

X_{7it} = Debt-to-Equity Ratio (DER) of company i at period t

X_{8it} = Earning Yield (EY) of the company i at period t

$X_{9t,}$ D_t = The Rusia-Ukraine War at period t

a = the intercept of the regression model

$\beta.$ = the slope coefficient

εit = the error component of the observed cross-sectional units and period

Analysis with regression of panel data has several advantages, among others, namely (Gujarati, 2003):

1. Panel data can take into account individual heterogeneity explicitly by allowing individual-specific variables.
2. This ability to control individual heterogeneity further makes panel data used to test and build more complex behavioural models.
3. Panel data is based on repeated cross-section observations (time-series), so the panel data method is suitable for studying dynamic adjustment.
4. The high number of observations impacts more informative, more varied, and collinearity data between diminishing variables and increased degrees of freedom (degrees of freedom-df) to obtain more efficient estimates.
5. Panel data can be used to study complex behavioural models.
6. Panel data can minimize the possible aggregation of individual data.

3.4.3 Panel Data Model Approaches

Referring to the advantages mentioned above, then in the panel data model there should be no testing approach in regressing panel data consisting of pooled least square (common effect), fixed effect approach (fixed effect), and random effect approach (random effect).

3.4.3.1 Common Effect Model (CEM)

Common effect model (CEM) or pooled least square (PLS) is a model obtained by combining or collecting all cross-data and time-guided data (Gujarati, 2016). This data model is then estimated using ordinary least square (OLS), which is the simplest method of linear regression (Baltagi, 2005). The Common-Effect Model equation is as follows:

$Yit = \alpha + \beta Xit + eit$ (3.14)

I = 1, …., n

t = 1, …., T

where,

Yit = the dependent variable of the cross-sectional units over the time period observed

Xit = the independent variable of the cross-sectional units over the time period observed

α = the intercept of the regression model

β = the slope coefficient

eit = the error component of the observed cross-sectional units and time period

n = the number of observed cross-sectional units

T = the number of observed time period

3.4.3.2 Fixed Effect Model (FEM)

Fixed effect model (FEM) can solve the problem of interception assumptions or slopes from regression equations that are considered constant in

the pooled least square model (Gujarati, 2016). This method uses dummy variables to generate different parameter values across cross-section units and between time (time-series). The Fixed-Effect Model equation is as follows:

$$Yit = \beta Xit + ai + eit \quad (3.15)$$

I = 1, …., n

t = 1, …., T

where,

Yit = the dependent variable of the cross-sectional units over the time periodobserved

Xit = the independent variable of the cross-sectional units over the time periodobserved

ai = the regression model intercept of the observed cross-sectional units and/orthe observed time period

β = slope coefficient

eit = the error component of the observed cross-sectional units and time period

n = the number of observed cross-sectional units

T = the number of observed time period

3.4.3.3 Random Effect Model (REM)

The random-effect model (REM) is used to estimate panel data where interference variables may be interconnected between time and between individuals (Gujarati, 2016). In this model, different parameters between time and between individuals are entered into errors because this model is often also referred to as the Error Component Model (ECM). The Random Effect Model equation is as follows:

$$Yit = \beta Xit + a + wit \quad (3.16)$$

$wit = \varepsilon it + eit$

I = 1, …., n

t = 1, …., T

Where :

Yit	= the dependent variable of the cross-sectional units over the time period observed
Xit	= the independent variable of the cross-sectional units over the time period observed
a	= intercept of the regression model
β	= slope coefficient
wit	= combination of two error components, namely εit and eit
εit	= individual-specific error component
eit	= the error component of the observed research sample and time period
n	= the number of observed cross-sectional units
T	= the number of observed time period

3.4.3.4 Selection of Panel Data Estimation Model

To determine the most suitable panel data model approach among the Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM), it's important to test the model's appropriateness (Gujarati, 2016). According to Wooldridge (2016), if the data does not include dummy variables, as is the case with the Fixed Effect Model, and control variables are included among the explanatory variables, the Random Effect Model (REM) is recommended. This view aligns with Nachrowi & Usman's (2016) guidance, which suggests that if the time period (T) is longer than the number of individuals (N), the Fixed Effect Model (FEM) should be used. Conversely, if the time period (T) is shorter than the number of individuals (N), the Random Effect Model (REM) is preferable.

To select the best panel data regression model, three tests can be used to assess model suitability: the Chow test, the Hausman test, and the Lagrange Multiplier test. Here's a brief explanation of each test:

1. **Chow Test**: Used to determine whether a Fixed Effect Model is more appropriate than a Common Effect Model by testing for the presence of individual-specific effects.
2. **Hausman Test**: Helps in choosing between the Fixed Effect Model and the Random Effect Model by checking if the unique errors are correlated with the regressors, indicating whether the Random Effect Model is appropriate.
3. **Lagrange Multiplier Test**: Used to decide between a Random Effect Model and a Common Effect Model, testing whether the individual effects are significantly different from zero.

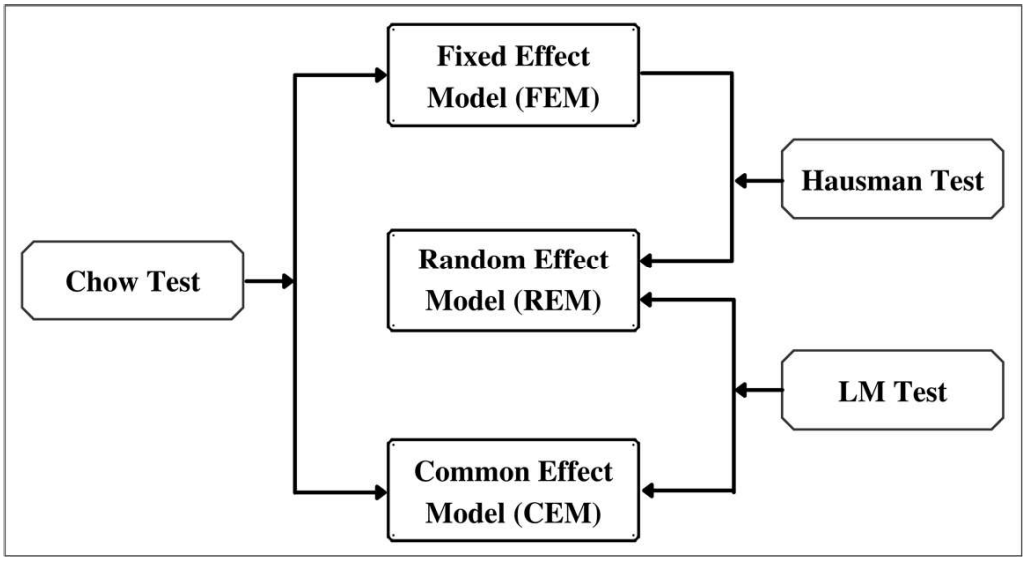

Figure 3.2 Fit Test Model

1. Chow Test

The Chow Test (Chow Test) or restricted F test (Gujarati, 2003) is used to determine which model can best be used to estimate panel data, whether fixed effect model (FEM) or common effect model (CEM). The formula for obtaining statistical F values as formulated by Chow is as follows:

$$F = \frac{(PRSS - URSS)/(N-1)}{(URSS)/(NT-N-K)} \tag{3.17}$$

Where:

PRSS = is Residual Sum of Square (CEM)

URSS = is Residual Sum of Square (FEM)

N = is the amount of cross-section data

T = is the amount of time series data

k = is the number of free variables The null hypothesis of the restricted

F test is as follows:

Ho = Common effect model (CEM) is better than fixed-effect model (FEM)

Ha = Fixed effect model (FEM) is better than common effect model (CEM)

The hypothesis testing criteria is that if the F count value results > F table at a certain level of α confidence, then Ho is rejected, H1 is accepted, meaning the fixed effect model is more appropriately used for estimation techniques (Aulia, 2004).

2. Hausman Test

The Hausman test is used to choose which model is best used to estimate panel data, whether fixed effect model (FEM) or random effect model (REM) (Gujarati, 2003). The formula for obtaining Hausman's test scores is as follows:

$$m = (\beta - b)\,(M0 - M1)\text{-}1(\beta - b) \approx X2\,(K) \qquad (3.18)$$

Where:

β = is a vector for fixed effect variable statistics

b = is a vector for random effect variable statistics

M0 = is a covariance matrix for fem conjecture

M1 = is the covariance matrix for alleged REM Hypothesis zero of the Hausman test is as follows:

Ho = Random effect model (REM) is better than fixed-effect model (FEM)

H1 = Fixed effect model (FEM) is better than common effect model (CEM)

The hypothesis testing criteria if X2 calculates > X2 tables and the p- value is significant, then Ho is rejected, and the fixed effect model is appropriate for use. (Gujarati, 2016).

3. Lagrange Multiplier Test

The Lagrange Multiplier test or LM Test is used to choose which model is best used to estimate panel data, whether random effect model (REM) or common effect model (CEM) (Gujarati, 2016). The formula for obtaining the Lagrange multiplier test value is as follows:

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}(\sum_{t=1}^{T} e_{it})^2}{\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}^2} - 1\right]^2 \quad (3.19)$$

Where:

N : Number of individuals Q : Number of periods

e : Residual from model OLS

The null hypothesis of the LM Test is:

Ho = Common effect model (CEM) is better than random effect model (REM)

H1 = Random effect model (REM) is better than common effect model (CEM)

With the hypothesis testing criteria, if X^2 calculates > X^2 table and the p-value is significant, Ho is rejected, meaning the REM model is more appropriately used (Gujarati, 2016).

3.4.4 Classical Assumption Test

A classical assumption test is a statistical method used to assess the relationship between variables. It includes tests for linearity, normality, autocorrelation, multicollinearity, and heteroscedasticity (Basuki & Yuliadi, 2015). However, Iqbal (2015) argues that not all of these tests are applicable when using

panel data. Specifically, the linearity test is often omitted as linear regression models are typically assumed to be linear. The normality test may not be suitable for studies involving multiple companies, panel data, or secondary data. The autocorrelation test is more relevant for time-series data, while the heteroscedasticity test is not necessary for panel data since heteroscedasticity is usually expected. Consequently, the multicollinearity test is the primary test recommended for this type of study because it involves more than two independent variables. Additionally, the author conducted a normality test, assuming that coal mining companies have similar business models.

The multicollinearity test identifies the presence of high correlations among variables in a regression model (Ainiyah et al., 2016). This can be assessed using the Variance Inflation Factor (VIF). If the VIF is below 10, there is no multicollinearity issue among the independent variables (Hair Jr et al., 2010). Conversely, a VIF of 10 or higher indicates a multicollinearity problem. While multicollinearity does not significantly impact the ability of a regression equation to predict the dependent variable, it poses a concern when assessing the relative importance of independent variables with high VIFs. To address multicollinearity, researchers can replace the dependent variable without altering the independent variables, combine cross-sectional and time-series data, or increase the sample size (Basuki & Prawoto, 2016).

3.4.5 Significant Test

Hypothesis testing aims to determine the influence of independent variables (X) with dependent variables (Y) using linear regression of panel data. The steps of this analysis are as follows:

3.4.5.1 Partial Test (t-test)

The t-test shows how far the influence of one individually independent variable affects explaining the dependent variable's variation. The t-test is used to test the regression coefficient of its independent variable partially. The procedures

used to perform the t-test are:

- Formulating a hypothesis
- H_i; $\beta_1=\beta_2=\beta_3 \neq 0$, means that the independent variable has a significant effect on the dependent variable partially.
- Determining the level of significance

  This hypothesis was tested using a significance level of = α

  Were,

  α = Highly Significant: p-value < 0.01

  α = Significant: 0.01< p-value < 0.05

  α = Marginally Significant: < 0.05 p-value < 0.1
- Determine the research hypothesis testing criteria:
- Based on the comparison of t_{count} with t_{table} with guidelines:

  3.3.1.1.1 If $t_{count}<t_{table}$ means that the independent variable is partially significantand does not significantly affect the dependent variable.

  3.3.1.1.2 If $t_{count}>t_{table}$ means that the independent variable partially has asignificant influence on the dependent variable.
- Based on the p-value, the conditions are:
  1. If the p-value > α, the independent variable partially does not significantly affect the dependent variable.
  2. If the p-value < α, the independent variable partially has a significant influence on the dependent variable.

3.4.5.2 Simultaneous Significance Test (F-Test)

A simultaneous significance test (F-test) is used to determine whether all independent variables have the same effect on the dependent variable. The hypotheses of the F-test are as follows:

H0: All parameters = 0

Ha: At least one parameter ≠ 0

Two criteria are used to measure the hypotheses as mentioned above:

a. If p-value < significance level of 0.05, the H0 is rejected = all of the

independent variables have the same effect on the dependent variable

b. If p-value > significance level of 0.05, the H0 is not rejected = all of the independent variables do not have the same effect on the dependent variable

3.4.5.3 Coefficient of Determination Test (Adjusted R^2)

The coefficient of determination (R^2) essentially measures how far the model's ability to explain the dependent variable is. The coefficient of determination value is between zero and one. A small value of R2 means that the ability of the independent variables in explaining the variation of the dependent variable is very limited. A value close to one means that the independent variables provide almost all the information needed to predict the variation of the dependent variable. In general, the coefficient of determination for cross-sectional data is relatively low due to the large variation between observation, while for time series data, it usually has a high coefficient of determination.

One thing to note is the spurious regression problem. (Ghozali, 2013) emphasizes that the coefficient of determination is only one and not the only criteria for choosing a good model. The reason is that if a linear regression estimate produces a high coefficient of determination but is not consistent with the economic theory of high determination but is inconsistent with the economic theory chosen by the researcher, or does not pass the classical assumption test, then the model is not a good estimator model. and should not be selected as an empirical model.

The basic weakness of using the coefficient of determination is the bias towards the number of independent variables included in the model. Every additional one independent variable, then R2 must increase no matter whether the variable has a significant effect on the dependent variable. Therefore, many researchers recommend using the adjusted R2 value when evaluating which regression model is the best. Unlike R2, the value of adjusted R2 can fluctuate if one independent variable is added to the model.

In fact, the adjusted R2 value can be negative, although what is desired must

be positive. According to Gujarti (2003) and Imam Ghozali (2013) if in the empirical test the adjusted R2 value is negative, then the adjusted R2 value is considered to be zero. Mathematically if the value of R2 = 1, then Adjusted R2 = R2 = 1 while the value of R2 = 0, then adjusted R2 = (1 = k)/(n – k) if k > 1, then adjusted R2 will be negative Ghozali (2013).

3.5 Research Hypothesis Testing

- Hypothesis 1:
  - Coal Price

    H_{01}: β_1=0

    H_{a1}: β_1>0
- Hypothesis 2:
  - Exchange Rate

    H_{02}: β_1=0

    H_{a2}: $\beta_1 > 0$
- Hypothesis 3:
  - Market Return

    H_{03}: β_1=0

    H_{a3}: $\beta_1 > 0$
- Hypothesis 4:
  - Return on Equity (ROE)

    H_{04}: β_1=0

    H_{a4}: $\beta_1 > 0$
- Hypothesis 5:
  - Current Ratio (CR)

    H_{05}: β_1=0

    H_{a5}: $\beta_1 > 0$
- Hypothesis 6:
  - Total Asset Turnover (TATO)

    H_{06}: β_1=0

H_{a6}: $\beta_1 > 0$

- Hypothesis 7:
  - Debt-to-Equity Ratio (DER)

    H_{07}: $\beta_1=0$

    H_{a7}: $\beta_1 > 0$
- Hypothesis 8:
  - Earning Yield (EY)

    H_{08}: $\beta_1=0$

    H_{a8}: $\beta_1 > 0$
- Hypothesis 9:
  - The Rusia-Ukraine War

    H_{09}: $\beta_1=0$

    H_{a9}: $\beta_1 > 0$
- Hypothesis 10:
  - Coal Price * The Rusia-Ukraine War

    H_{010}: $\beta_1=0$

    H_{a10}: $\beta_1>0$
- Hypothesis 11:
  - Exchange Rate* The Rusia-Ukraine War

    H_{011}: $\beta_1=0$

    $Ha11$:$\beta1>0$
- Hypothesis 12:
  - Market Return* The Rusia-Ukraine War

    H_{012}: $\beta_1=0$

    H_{a12}: $\beta_1>0$
- Hypothesis 13:
  - Return on Equity (ROE)* The Rusia-Ukraine War

    H_{013}: $\beta_1=0$

    H_{a13}: $\beta_1>0$

- Hypothesis 14:
  - Current Ratio (CR)* The Rusia-Ukraine War

    H_{014}: β_1=0

    H_{a14}: β_1>0
- Hypothesis 15:
  - Total Asset Turnover (TATO)* The Rusia-Ukraine War

    H_{015}: β_1=0

    H_{a15}: β_1>0
- Hypothesis 16:
  - Debt-to-Equity Ratio (DER)* The Rusia-Ukraine War

    H_{016}: β_1=0

    H_{a16}: β_1>0
- Hypothesis 17:
  - Earning Yield (EY)* The Rusia-Ukraine War

    H_{017}: β_1=0

    $Ha17$:$\beta1$>0

3.6 Research Process

Figure 3.2 below shows the research process. It can be seen that the first step of the research process is to obtain descriptive statistical results consisting of mean, standard deviation, and correlation coefficient. The second step of the research process is to run an average difference test consisting of an independent sample t-test and a paired sample t-test. The third step of the research process is to identify the free variable (X) and the bound variable (Y). Once the free variable (X) and the bound variable (Y) are identified, the fourth step of the research process is to form a data panel to be tested to select the most suitable panel data model between common-effect, fixed-effect, and random-effect models for regression. In choosing the panel data model best suited for regression, the fifth step of the research process is to perform the Chow Test, Hausman Test, and Lagrange Multiplier Test. The sixth stage of the research process is to perform theclassical

assumption test of multicollinearity test against multiple linear regressionequations. The seventh stage of the research process is to conduct a significant test consisting of the Individual Regression Coefficient Test (Test t), F-Test, and the Coefficient of Determination Test (Adjusted R2), against multiple linear regression equations. Step eight, or the end of the research process, is to interpret the test results.

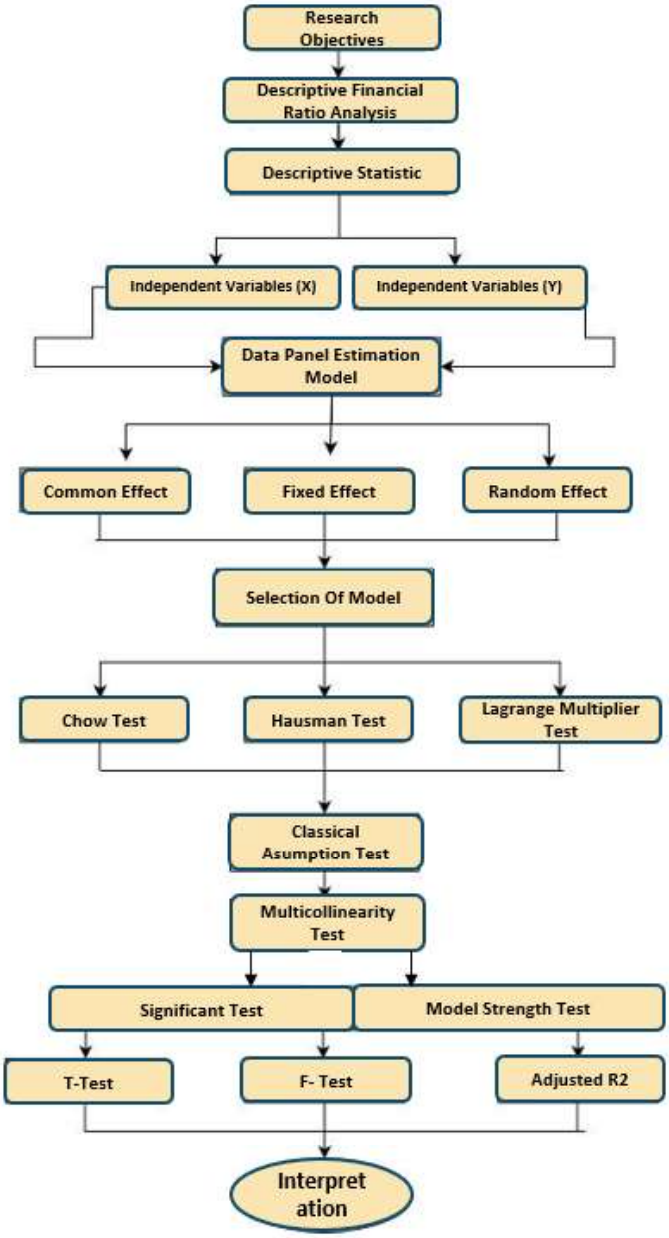

Figure 3.3 Research Process