CHAPTER 3

METHODOLOGY

1.1 Research Design

Bougie and Sekaran (2017) explain research design as a plan to collect, measure, and analyze data based on research questions. This study uses a quantitative research design. Refer to Sugiyono (2017), quantitative methods are carried out because of numbers and data analysis. Quantitative research emphasizes quantifying variables and using statistical analysis to test the hypothesis. The key method is correlational research design, which examines the relationship between two or more variables without manipulating them. Figure 3.1 shows each research step, which is comprised of identifying the research topic, literature review, and problem definition, developing the research question, constructing a research design, data collection and pre-test, data analysis, conclusion, and recommendation.

Figure 3.1 Research Design

The design used a quantitative approach with purposive sampling to acquire the necessary sample data. The method aims to investigate quantitative features, phenomena, and their interactions scientifically. Winter (2000) found that quantitative research improves validity by using statistical and mathematical approaches to measure results conclusively. The quantitative analysis tests hypotheses suggested in the study using various statistical tools to identify correlations among variables.

3.2 Research Procedure

This chapter explains the statistical procedures utilized in the study, which are compatible with earlier work. The research aims to evaluate the consumer cyclical industry's financial performance by analyzing data and examining the relationship between the dependent variable (the consumer cyclical industry's stock return and profitability) and multiple independent variables (fundamental and macroeconomic variables), as well as a moderating factor (a dummy variable from the Covid-19 period) to determine any impact on the variables.

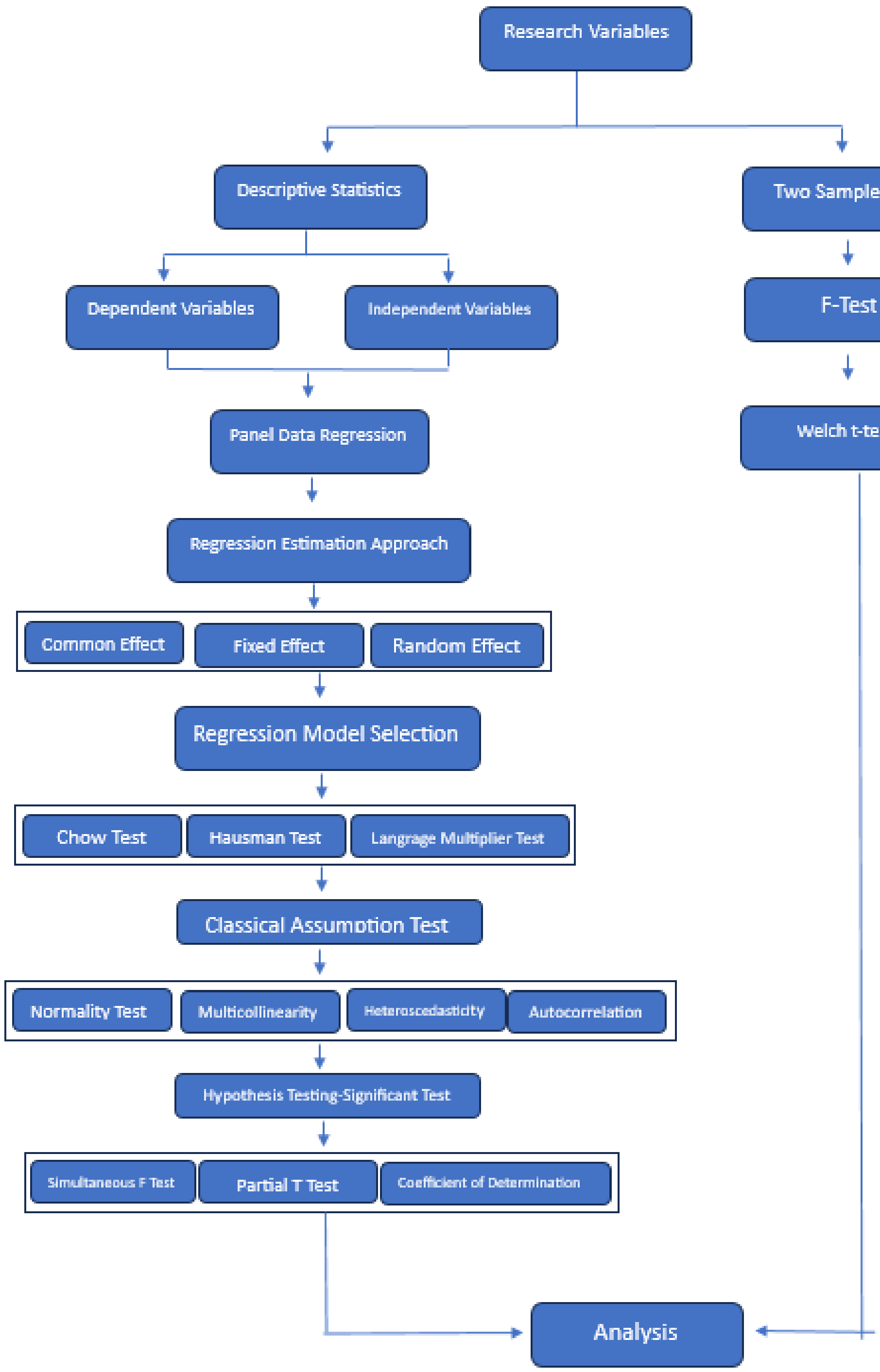

Figure 3.2 Research Procedure

Figure 3.2 shows the research procedure in this study. The research began with descriptive statistics after collecting all necessary data variables to comprehend the data and its aspects better. Then, data were analyzed using panel data regression to analyze the factors affecting consumer cyclical stock performance.

The regression model for consumer cyclical stock performance included a dummy variable, COVID-19, to account for the pandemic's impact on macroeconomic indicators and stock returns. When selecting the regression model, the variables of each regression model were examined to see whether any of the OLS Assumptions were violated.

Classic assumptions testing included normality, multicollinearity, autocorrelation, and heteroscedasticity assessment. After selecting a regression model, a significance test was used to evaluate the relationship between the independent and dependent variables. The significance test results were evaluated to determine the effect of independent factors.

The second section of the analysis compared consumer cyclical stock performance before and during the epidemic timeframe. The Welch t-test was used on two-sample data to identify potential differences in means. The F-test Two Sample for Variances was used to assess variations in stock returns before and during the COVID-19 pandemic. The assessment aimed to investigate the impact of the COVID-19 epidemic on consumer cyclical stock returns and variance performance.

The investigation concluded with an analysis of consumer cyclical 's financial performance. Financial data from consumer cyclical statements were used to calculate financial ratios. All test data were examined for insights and evidence.

3.3 Data Collection

This study uses secondary data collection. According to Levine, et al., (2014), secondary data sources refer to data acquired by others for analysis purposes. The data used to conduct the study was collected from Gurufocus.com and the Capital Market Electronic Library at the Indonesia Stock Exchange (www.idx.com). All the data is collected from the financial report quarterly for the period of 10 years from the year 2014-2023, The data for this study were collected from the IDX categories of consumer cyclical companies which have ESG Leader Index (7 Company) and have ESG Score (3 companies) and active listed in IDX in 2024, the total population of the companies selected is ten companies.

3.4 Research Population and Sample

According to "Levine, et al. (2014)", a sample is a subset of a population chosen for analysis. The sample analysis yields estimates for the complete population's characteristics. This study uses a non-probability sampling strategy. In a nonprobability sample, we choose things or individuals without knowing their likelihood of selection (Levine, et al., 2014).

A nonprobability sample can be divided into two subcategories: convenience samples and judgment samples. In this situation, the judgment sample is selected as a subcategory of the nonprobability sample. Judgment sampling occurs when the elements picked for the sample are chosen by the researcher's judgment.

This study sampled companies that met the parameters listed below.

1. The company is categorized as consumer cyclical in IDX
2. The company has an ESG score or is categorized as an ESG Leader Stock Index.
3. The company still exists in 2024.
4. The data being analyzed is the quarterly data.
5. Data used from the quarterly report is for the period 2014-2023.

Table 3.1 Table Population and Sample

Description	Population	Sample
Object	Consumer Cyclical Company registered in IDX period 2014-2023	Consumer Cyclical Company registered in IDX period 2014-2023 and become ESG Leader Index or have ESG Score

Numbers of Object	166 companies	ten companies (7 companies listed in 2014-2023,

Table 3.2 shows the list of the companies from the consumer cyclical industry that will be used for analysis.

Table 3.2 List of Sample

No.	Code	Company Name	ESG Score / ESGL Stock
1	MAPI	Mitra Adiperkasa Tbk	√
2	AUTO	Astra Otoparts Tbk	√
3	BOGA	PT Bintang Oto Global Tbk	√
4	ERAA	Erajaya Swasembada Tbk	√
5	FILM	PT MD Pictures Tbk.	√
6	GJTL	Gajah Tunggal Tbk	√
7	MAPA	Map Aktif Adiperkasa Tbk	√
8	MPMX	PT Mitra Pinasthika Mustika Tbk.	√
9	SCMA	Surya Citra Media Tbk	√
10	MNCN	Media Nusantara Citra Tbk	√

3.5 Econometric Model and Variable

Table 3.3 Table Dependent and Independent Variables

Dependent Variables	Variables description	Source
$\Delta\text{Stock Return}_t = \frac{P_t}{Pt - P_{t-1}}$ (3.1)	Indicator of return of a stock. Pt is the close price at the period (eg. period t) and Pt-1 is the close price one period before period t.	Thamrin and Sembel (2020) Widjaja, et al. (2023) Chiang, et al. (2024)
Independent Variables	**Variables description**	**Source**
Asset Efficiency = Total Asset Turnover = $TAT_t = \frac{\text{Net Sales}_t}{\text{Average Total Assets}_t}$ (3.2)	Indicator of how efficient the company's operations are in leveraging its assets to create income.	Munawar (2019) Thamrin and Sembel (2020)
Liquidity = Current Ratio = $CR_t = \frac{\text{Current Assets}_t}{\text{Current Liabilities}_t}$ (3.3)	Indicator of the company's capacity to cover its short-term liabilities with its short-term assets.	Adawiyah and Setiyawati (2019) Widjaja, et al. (2023) Thamrin and Sembel (2020)
Profitability = Return on Equity = $ROE_t = \frac{\text{Net Income}_t}{\text{Equity}_t}$ (3.4)	Measurement of return earned by common shareholders investing in the firm.	Malau (2020) Ihsan, et al. (2023) Widjaja, et al. (2023)
Market Return = ΔMR_t =	Measurement of change in stock market return. Pt is the close	Thamrin and Sembel (2020)

$\frac{JKSE_t - JKSE_{t-1}}{JKSE_{t-1}}$ (3.5)	price at the period (eg. period t) and Pt-1 is the close price one period before period t.	Ihsan, et al. (2023) Chiang, et al. (2024)
Exchange Rate = log (ER_t) (3.6)	Measurement of the foreign exchange rate at the period.	Thamrin and Sembel (2020) Ihsan, et al. (2023)
Interest rate =log (IR_t) (3.7)	Measurement of interest rate at the period.	Thamrin and Sembel (2020) Mensah et al. (2020)

3.6 Statistical Method

3.6.1 Panel Data Regression Analysis

Panel data analysis combines the strengths of time series with cross-sectional data, allowing for a better understanding of complex processes.

In research, panel data are analyzed using regression (Xu et al., 2007).

The first regression equation model tests fundamental financial factors that affect consumer cyclical stock return:

STOCK_RETURN = C(1) + C(2)*ASSET_EFFICIENCY + C(3)*LIQUIDITY + C(4)*ROE + C(5)*MARKET_RETURN+C(6)*EXCH_RATE+C(7)*INTEREST_RATE (3.8)

The second regression equation model tests macroeconomic factors that affect consumer cyclical stock return:

STOCK_RETURN = C(1) + C(2)*EXCH_RATE + C(3)*INTEREST_RATE + C(4)*MARKET_RETURN + C(5)*EXCXCOVID + C(6)*INTRSTXCOVID + C(7)*MARKTXCOVID

(3.9)

Panel data regression analysis has several benefits (Gujarati, 2004):

1.Panel data can account for individual heterogeneity by including person-specific characteristics.

2.Panel data's capacity to control individual heterogeneity allows for the testing and development of complicated behavioral models.

3.Panel data, which uses repeated cross-section observations (time series), is ideal for analyzing dynamic adjustments.

4. Having a large number of observations leads to more informative, diversified, and collinear data between variables, and greater degrees of freedom (df) for more effective estimation.

5. Panel data can be utilized to investigate more complicated behavioral models.

6. Panel data can reduce the potential aggregation of individual data.

3.6.2 Panel Data Model Approaches

Panel data analysis comprises three more or less independent methodologies.

3.6.2.1 Common Effect Model (CEM)

The Common Effect Model (CEM) or Pooled Least Squares (PLS) is a model that combines cross-data and time-guided data (Gujarati, 2016). The data model is calculated using the simplest linear regression approach, ordinary least squares (OLS). (Baltagi, 2008).

3.6.2.2 Fixed Effect Model

The fixed effect model (FEM) addresses the issue of interception assumptions or slopes from regression equations that are constant in the pooling least square model

(Gujarati, 2016). This method uses dummy variables to generate varying parameter values across cross-section units and across time (time series).

3.6.2.3 Random Effect Model

The random-effect model (REM) is used to estimate panel data with interconnected variables over time and across individuals (Gujarati, 2016). This model, also known as the Error Component Model (ECM), incorporates individual and temporal factors into errors.

3.6.3 Selection of Panel Data Estimation Model

Determine the optimal technique for the three-panel data model and assess its applicability. (Gujarati, 2016).

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 null hypothesis of the restricted F test is as follows:

H0 = Common effect model (CEM) is better than the fixed-effect model (FEM)

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

According to Aulia (2004), if the F count value exceeds the F table at a specific degree of confidence, H0 is rejected and H1 is accepted, indicating that the fixed effect model is better suited for estimating.

2. Hausman Test

The Hausman test determines which model is optimal for estimating panel data: fixed effect model (FEM) or random effect model (REM) (Gujarati, 2003).

Hypothesis zero of the Hausman test is as follows:

H0 = Random effect model (REM) is better than fixed-effect model (FEM)

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

Hypothesis testing criteria: If X2 calculates > X2 tables and the p-value is significant, H0 is rejected and the fixed effect model is suitable for usage (Gujarati, 2016).

3.Lagrange Multiplier Test

The Lagrange Multiplier Test (LM Test) determines which model is appropriate for estimating panel data: random effect model (REM) or common effect model (CEM) (Gujarati, 2016)

H0 = Common effect model (CEM) is better than random effect model (REM)

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

If X^2 calculates > X^2 table and the p-value is significant, H0 is rejected, indicating that the REM model should be employed instead (Gujarati, 2016).

3.6.4 Classical Assumption Test

Classic Assumptions Testing aimed to assess the robustness of the obtained regression model and whether the model generated statistical estimates that were unbiased. The tests consist of the linearity test, normality test, autocorrelation test, and multicollinearity test (Gujarati and Damodar, 2004). The tests would assess if the suggested regression model met the key Ordinary Least Squares (OLS) assumption.

Linearity Test

The first test was the linearity test, which assessed the parameters in the regression model. The linearity test was satisfied when the dependent variable was a linear function of the selected independent variables plus the error term. (Fox, 2015).

Normality Test

The normality test determines if the dependent and independent variables in a regression model have normal distributions (Burton, 2021). A regression model is

considered robust if the data variables are regularly distributed or nearly so.
A Normal QQ plot of residuals can be used to visually evaluate the normality assumption. A normal probability plot, also known as a quantile-quantile (Q-Q) plot, shows the data distribution compared to the expected normal distribution. Observations in properly distributed data should follow a linear pattern. If the data is not regularly distributed, the points may create a curve that deviates from the straight line.

Multicollinearity Test

Multicollinearity refers to an ideal or near-perfect linear relationship between some or all independent variables in a regression model. Ainiyah et al. (2016) employed the multicollinearity test to identify strong correlations between variables in a regression model.

Multicollinearity testing can be performed by calculating the Variance Inflation Factors (VIF) of the variables. Hair et al. (2019) found that a Variance Inflation Factor (VIF) of less than ten indicated a low likelihood of multicollinearity among independent variables. If the Variance Inflation Factors (VIF) were equal to or greater than 10, it was likely that the independent variables were highly associated. The regression equation's ability to predict the dependent variable remained unaffected by multicollinearity.

According to Shrestha (2020), a correlation coefficient method can detect multicollinearity.

Heteroscedasticity Test

Heteroscedasticity is a statistical term defined as unequal variances in a regression model or unequal scatter of residuals or error terms, this is the condition in which the variance of the error terms in a regression model is not constant. There are several methods to detect heteroscedasticity:

▪Visual method: plotting the residuals against the independent(s) variable sometimes can reveal a non-random pattern, suggesting unequal variance.

▪Formal tests: statistical tests such as the Breush-Pagan Test or White Test can formally assess heteroscedasticity.

Auto-correlation Test

The autocorrelation test was used to determine whether there was a series connection between the variables and error terms (Burton, 2021). According to the no autocorrelation OLS assumption, the error terms of different observations should be uncorrelated. To provide reliable estimates, the regression model's variance should be constant, with no heteroscedastic errors. The OLS assumption implies that the error terms are independent and identically distributed (IID). There are several ways to test autocorrelation. A study by Novia (2012) showed that the accuracy of testing autocorrelation with the Breusch-Godfrey method is closer to testing the existence of autocorrelation than the Durbin-Watson method.

Hypothesis:

H0: There was no self-correlation between errors.

H1: There was self-correlation between errors.

3.6.5 Partial t-test

The t-test evaluated the impact of individual independent variables on the dependent variable by analyzing their regression coefficients. Hypothesis statements evaluated the importance of regression coefficients:

$$H0: \beta j = 0 \tag{3.10}$$

$$H1: \beta j \neq 0$$

The test statistic for this t-test was derived from the t distribution, using the t-statistic formula:

$$t = \frac{\widehat{\beta}_j}{SE(\widehat{\beta_j})} \tag{3.11}$$

•If the p-value was less than 0.05, then the null hypothesis was rejected. This provided substantial evidence that the null hypothesis was erroneous and the alternative hypothesis was adopted. The independent variable had a significant impact on the dependent variable.

•If the p-value was more than 0.05, the null hypothesis could not be rejected. This suggested that there was no substantial evidence to support the alternate hypothesis, and the effect proposed in the alternative hypothesis, hence, lacked statistical significance. The independent variable did not significantly impact the dependent variable.

3.6.6 F-test Two Sample for Variances

The F test evaluates whether two independent samples from two populations have identical variances. A two-tailed version evaluates the alternative hypothesis that the variances are not equal. A one-tailed test determines if the variance from the first population is larger or less than that of the second population.

The formula for comparing two variances with the f-test is:

$$F = \frac{s_1^2}{s_2^2} \tag{3.12}$$

Where:

F = F value
s_1^2 = sample 1 variance
s_2^2 = sample 2 variance.

Hypothesis

$H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$ (for a two-tailed test)

$\sigma_1^2 < \sigma_2^2$ (for a one-tailed test)

$\sigma_1^2 > \sigma_2^2$ (for a one-tailed test) (3.13)

3.6.7 Welch Test

The Welch t-test is often used to compare two samples from two populations, regardless of their variances. The Welch t-test differs from the standard Student t-test by not assuming equal variances amongst samples. The Welch t-test is widely used to compare two samples, as real-world data may not have comparable variances. The t-test determines if the sample means of two independent populations are equal.

The formula for computing the Welch t-test:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{SE_1^2 + SE_2^2}} \quad (3.14)$$

Using null and alternative, as follows:

$H_0: \mu_1 = \mu_2$ the population means are equal

$H_1: \mu_1 \neq \mu_2$ the population means are not equal