

AUDITING | RESEARCH ARTICLE

The Relationship Between the Beneish M-Score and the Altman Z-Score: Evidence from Energy Sector Companies in Indonesia

Agung Rachmad Prasetyo¹, Linda Santioso²

^{1,2}Department of Accounting, Faculty of Economics and Business, Universitas Tarumanagara, Jakarta, Indonesia.
Email: agung.125249404@stu.untar.ac.id¹

ARTICLE HISTORY

Received: May 26, 2026
Revised: June 17, 2026
Accepted: June 30, 2026

DOI

<https://doi.org/10.52970/grar.v6i2.2405>

ABSTRACT

The energy sector plays a strategic role in the Indonesian economy; however, its high exposure to commodity price volatility makes it particularly vulnerable to irregularities in financial reporting. This study examines the relationship between the Beneish M-Score and the Altman Z-Score in detecting fraudulent financial reporting among energy companies listed on the Indonesia Stock Exchange (IDX) during the 2022–2024 period. Employing a quantitative descriptive-comparative research design, this study analyzes 183 firm-year observations from 61 companies selected through purposive sampling. The analysis consists of three stages: classification based on the threshold values of each model; cross-tabulation combined with the chi-square test and Cramér's V to evaluate the consistency between the two classification models; and pooled ordinary least squares (Pooled OLS) regression with robust standard errors to examine whether financial distress influences earnings manipulation. The findings indicate that 38.80% of the observations are classified as potential manipulators according to the Beneish M-score, while 42.62% are categorized as financially distressed based on the Altman Z-score. A statistically significant association was found between the two classification models ($\chi^2 = 11.83$, $p = 0.003$; Cramér's $V = 0.254$). However, the distribution exhibited an unexpected U-shaped pattern, with the highest proportion of potential manipulators observed in the safe zone (52.24%), followed by the distress zone (37.18%), and the lowest proportion in the grey zone (18.42%). Consistent with this non-monotonic relationship, neither the overall Altman Z-score nor any of its five financial ratio components had a statistically significant effect on the Beneish M-score in the linear regression model. These findings suggest that financial distress alone is insufficient to explain earnings manipulation in the energy sector and highlight the importance of monitoring financially healthy firms, particularly during commodity price booms.

Keywords: Beneish M-Score, Altman Z-Score, Fraudulent Financial Reporting, Financial Distress, Energy Sector.

JEL Code: G32, M41, M42

I. Introduction

Financial statements are the primary means through which corporate management communicates with shareholders, creditors, regulators and the broader market. Therefore, their reliability is not merely a matter of internal accounting practices but also of public interest. When financial statements no longer reflect



the underlying economic reality, capital may be misallocated, debt covenants may be misinterpreted, and market confidence may decline. The 2024 Report to the Nations published by the Association of Certified Fraud Examiners (ACFE) indicates that although financial statement fraud accounts for only 5% of occupational fraud cases, it results in a median loss of USD 766,000 per case, which is more than six times the median loss caused by asset misappropriation. The mining and energy industries are prominently represented in the report, with the mining sector recording the highest median loss per case among the 22 industries surveyed (ACFE 2024). In the Indonesian context, the energy sector contributes significantly to state revenue, with Non-Tax State Revenue generated from energy and mineral resources amounting to hundreds of trillions of rupiah annually. Indonesia is the world's largest exporter of thermal coal, supplying a substantial share of the global thermal coal demand. This scale is particularly important because any distortion in the financial statements of energy companies has the potential to affect investment decisions, royalty calculations, and tax revenues worth trillions of Rupiah. Several real-world cases have reinforced these concerns. PT Timah Tbk represents one of the most significant examples, with court findings reporting state losses of approximately Rp271 trillion resulting from illegal arrangements during 2015–2022, including the falsification of financial documents (Saputri & Irene, 2024). Likewise, PT Bumi Resources Tbk and PT Cakra Mineral Tbk are associated with asset overstatements and fictitious ownership of shares in subsidiaries (Christian et al., 2024; Marini et al., 2025). The period of 2022–2024 provides a particularly relevant setting for investigating these issues. The Russia–Ukraine conflict triggered a sharp increase in global commodity prices in early 2022, causing Indonesia's coal reference price to exceed USD 300 per ton by the end of that year, before declining dramatically to below USD 130 per ton in early 2024. This represents a decrease of nearly 62% within less than two years. Such extreme price fluctuations placed management under dual pressure to report financial performance that maintained investor confidence while simultaneously preserving short-term performance, despite rapidly changing economic conditions. According to the Fraud Triangle theory proposed by Cressey (1953), this type of financial pressure constitutes one of the primary conditions that encourage fraudulent financial reporting.

Two quantitative models are widely recognized in the literature for detecting financial statement fraud using financial ratios: the Beneish M-Score and the Altman Z-Score. The Beneish M-Score, developed by Beneish (1999), is specifically designed to identify earnings manipulation using eight accounting ratios that measure changes in receivables, gross margin, asset quality, sales growth, depreciation, operating expenses, leverage, and accrual. In contrast, the Altman Z-score, introduced by Altman (1968), evaluates a firm's likelihood of bankruptcy. Although the two models assess different dimensions of corporate financial performance, they are theoretically related to each other. Rosner (2003), in a study published in *Contemporary Accounting Research*, found that firms approaching bankruptcy reported significantly higher discretionary accruals than financially healthy firms, demonstrating patterns similar to those of firms sanctioned by the SEC for fraudulent financial reporting. Similarly, Valaskova et al. (2021) used a sample of 11,105 firms from the Visegrad countries and concluded that financially distressed firms are more likely to engage in earnings management. Despite this theoretical relationship, previous empirical studies have produced inconsistent findings regarding the comparative performances of these models. Islamiyati and Purnomo (2021) reported that the Altman Z-score achieved an accuracy rate of 64%, whereas the Beneish M-score achieved only 37% among Indonesian manufacturing companies. Similar conclusions were reached by Tanusdjaja and Kurniawan (2018) and Putra (2021), who found that the Altman Z-score performed better than the Beneish M-score. In contrast, Pertiwi et al. (2023) found that the Beneish M-score was more accurate in detecting financial statement manipulation among Indonesian non-financial state-owned enterprises. Likewise, MacCarthy (2017) demonstrated that the Beneish M-score successfully identified manipulation signals at Enron for five consecutive years before the company's collapse. Miharsi et al. (2024) argued that these inconsistent findings stem from structural differences between the two models. The Beneish M-Score directly measures the symptoms of earnings manipulation, whereas the Altman Z-Score captures financial distress, which often serves as the underlying motivation for manipulation. Consequently, the effectiveness of each model depends on the industry, the observation period, and the prevailing macroeconomic conditions.

Against this background, several research gaps have been identified. First, most comparative studies have focused on manufacturing companies, consumer goods firms, and cross-industry samples. Kamaluddin et al. (2024) examined Malaysian listed companies during 2014–2018, while Valaskova et al. (2021) investigated firms from Central Europe. Neither study specifically examined the energy sector, which possesses distinctive accounting characteristics, including reserve recognition, take-or-pay contracts, and the Domestic Market Obligation (DMO) mechanism. Second, the 2022–2024 observation period has received little attention in previous studies. This period is unique because it was characterized by a commodity supercycle and significant regulatory changes, including Government Regulation (PP) No. 26 of 2022 concerning revised royalty rates. Third, although Kamaluddin et al. (2024) examined the influence of the Altman Z-score and its components on the Beneish M-score in Malaysia, similar evidence has not yet been provided for Indonesian energy companies operating under simultaneous financial distress and commodity boom conditions. Fourth, previous studies conducted in Indonesia have not integrated cross-tabulation analysis, the chi-square test, and regression analysis within a single analytical framework. This study addresses these research gaps by examining 61 energy companies listed on the Indonesia Stock Exchange (IDX) from 2022 to 2024. Specifically, it pursues four objectives: (1) to describe the Beneish M-Score classification of the sample companies; (2) to describe their Altman Z-Score classification; (3) to examine the consistency between the two classifications using cross-tabulation, the chi-square test, and Cramér's V; and (4) to investigate whether financial distress, as measured by the Altman Z-Score and its component ratios, significantly influences earnings manipulation measured by the Beneish M-Score. This study makes two primary contributions to the literature. Empirically, it provides the first comparative evidence on the application of the Beneish M-Score and Altman Z-Score to Indonesia's energy sector during a commodity supercycle. Theoretically, it identifies a U-shaped relationship between financial distress and earnings manipulation, thereby challenging the conventional linear interpretation of the Fraud Triangle theory and suggesting the need for non-linear approaches in future research.

II. Literature Review and Hypothesis Development

2.1. Agency Theory

Agency Theory, introduced by Jensen and Meckling (1976), explains the contractual relationship between principals (owners) and agents (managers). The theory emphasizes that differences in interests and information asymmetry create incentives for managers to act opportunistically. Jiang (2023) argues that agency problems arise from two fundamental conditions: conflicts of interest and information asymmetry. When managers possess more comprehensive information about a firm's actual condition than shareholders, moral hazard becomes an inherent organizational risk. In the energy sector, this risk is intensified by the complexity of production-sharing contracts, cost recovery mechanisms, and long-term take-or-pay agreements, all of which provide managers with greater opportunities to obscure a firm's actual financial condition. Manipulating financial statements under such circumstances is a direct consequence of unresolved agency conflicts (Chenkiani & Prasetyo, 2023).

2.2. Fraud Triangle Theory

The Fraud Triangle Theory, proposed by Cressey (1953), states that three conditions must coexist for fraud to occur: pressure, opportunity, and rationalization. Pressure refers to the motivation to commit fraud, which typically arises from financial stress or the need to achieve performance targets. Opportunity reflects weaknesses in internal controls that enable fraudulent activities to remain undetected, whereas rationalization refers to the justification perpetrators use to legitimize their unethical behavior. Wolfe and Hermanson (2004) later extended this framework by introducing capability as a fourth element, forming the Fraud Diamond Theory. Among these elements, pressure is the most relevant to this study. Energy companies

experiencing commodity price volatility and increasing financial distress are more likely to face heightened financial pressures. This condition is represented by the Altman Z-score, while the Beneish M-score captures the accounting signals that may indicate earnings manipulation resulting from such pressure.

2.3. Beneish M-Score

The Beneish M-Score, developed by Beneish (1999), is a probit-based model designed to detect earnings manipulation likelihood. The model was developed using a sample of 74 firms identified by the SEC as financial statement manipulators and approximately 2,332 non-manipulating firms in the control group. It incorporates eight financial ratios that capture different dimensions of potential earnings manipulation: the Days Sales in Receivables Index (DSRI), Gross Margin Index (GMI), Asset Quality Index (AQI), Sales Growth Index (SGI), Depreciation Index (DEPI), Sales, General, and Administrative Expenses Index (SGAI), Leverage Index (LVGI), and Total Accruals to Total Assets (TATA). The aggregate M-score is calculated using the following equation:

$$M = -4.840 + 0.920 DSRI + 0.528 GMI + 0.404 AQI + 0.892 SGI + 0.115 DEPI - 0.172 SGAI - 0.327 LVGI + 4.679 TATA$$

Firms with an M-score greater than -2.22 are classified as potential manipulators, whereas firms with an M-score equal to or less than -2.22 are classified as non-manipulators. MacCarthy (2017) demonstrated the predictive capability of the model by identifying earnings manipulation signals at Enron five years before its collapse. Similarly, Pertiwi et al. (2023) reported that the Beneish M-score achieved an accuracy rate of 89.29% when applied to Indonesian state-owned enterprises. Nevertheless, other studies have reported insignificant results in certain industries (Putra, 2021; Suheni & Arif, 2020), indicating that the model's effectiveness depends on the industry and research context characteristics. Several studies conducted in Indonesia have demonstrated that the Beneish M-Score is effective in detecting earnings manipulation through annual changes in financial ratios. Tanusdjaja and Kurniawan (2018) found that the model significantly improved the detection of fraudulent financial reporting, whereas Pertiwi et al. (2023) reported that it achieved the highest level of accuracy among the fraud detection models examined. Furthermore, Miharsi et al. (2024) emphasized that the Beneish M-Score was specifically developed to identify earnings manipulation, which represents the core characteristic of fraudulent financial reporting.

2.4. Altman Z-Score

The Altman Z-score, developed by Altman (1968), is a bankruptcy prediction model based on Multiple Discriminant Analysis using a sample of 66 publicly listed manufacturing firms in the United States. The model combines five financial ratios, namely Working Capital to Total Assets (WCTA), Retained Earnings to Total Assets (RETA), Earnings Before Interest and Taxes to Total Assets (EBITA), Market Value of Equity to Total Liabilities (MVTL), and Net Sales to Total Assets (NSTA), into a single financial distress index. The model is expressed as follows

$$Z = 1.2 WCTA + 1.4 RETA + 3.3 EBITA + 0.6 MVTL + 1.0 NSTA$$

Firms with a Z-score above 2.99 are classified as financially healthy (safe zone), those with scores between 1.81 and 2.99 are categorized within the gray zone, and firms with scores below 1.81 are classified as financially distressed. In the context of fraud detection, the Altman Z-score serves as an indirect measure of the pressure element in the Fraud Triangle Theory. Rosner (2003) demonstrated that firms approaching bankruptcy report significantly higher abnormal accruals than financially healthy firms, supporting the use of the Z-score as an indicator of financial pressure that may precede earnings manipulation. Within the Fraud

Triangle framework, financial distress represents the pressure that motivates managers to manipulate financial statements to maintain access to external financing, comply with debt covenants, or preserve investor confidence (Cressey, 1953; Rosner, 2003). Furthermore, Islamiyati and Purnomo (2021) reported that the Altman Z-score achieved an accuracy rate of 64% in identifying fraud-related conditions among Indonesian manufacturing companies.

2.5. Fraudulent Financial Reporting

Fraudulent financial reporting refers to the intentional misstatement or omission of financial information in financial statements to mislead users of those statements (Handoko et al., 2022). According to Herianti et al. (2023), managers often engage in such practices to conceal poor financial performance and to protect their positions within the organization. The most common forms of fraudulent financial reporting include improper revenue recognition and asset overstatement (Dinasmara & Adiwibowo, 2020). In the energy sector, fraudulent reporting frequently involves excessive capitalization of exploration costs, overstatement of proven reserves, and delayed recognition of impairment losses.

2.6. Consistency Between the Two Models (H1)

According to the Fraud Triangle Theory, firms experiencing greater financial pressure are expected to have a higher likelihood of engaging in earnings manipulation. Kamaluddin et al. (2024) found a statistically significant association between the Altman Z-score and Beneish M-score classifications based on 3,242 firm-year observations from the Malaysian capital market. Similarly, Valaskova et al. (2021) report consistent findings using a sample of 11,105 firms from the Visegrad countries. These studies suggest that, although the Beneish M-Score and Altman Z-Score measure different aspects of corporate financial conditions, the classifications generated by the two models are systematically related. Based on these findings, the first hypothesis is formulated as follows:

2.7. Influence of Financial Distress on Earnings Manipulation (H2)

Rosner (2003) provides empirical evidence that firms approaching bankruptcy report significantly higher abnormal accruals than financially healthy firms, indicating a causal relationship between financial distress and earnings manipulation. Building on this argument, Kamaluddin et al. (2024) examined this relationship using two regression models. Model 1 employs the aggregate Altman Z-score as the sole predictor of the Beneish M-score, whereas Model 2 includes the five individual financial ratio components of the Altman Z-score as independent variables. Their findings showed that four of the five components, namely WCTA, RETA, EBITA, and NSTA, had a statistically significant influence on the Beneish M-score among Malaysian listed companies. This study adopts the same two-model framework to examine whether a similar relationship exists among Indonesian energy companies during the 2022–2024 commodity supercycle. Accordingly, the second hypothesis and its sub-hypotheses are formulated as follows: The research hypotheses are as follows:

H1 : There is a significant association between the Altman Z-Score classifications (distress, gray, and safe zones) and the Beneish M-Score classifications (manipulator and non-manipulator) among energy companies listed on the Indonesia Stock Exchange during the 2022–2024 period.

H2 : The aggregate Altman Z-Score has a significant effect on the Beneish M-Score among energy companies listed on the Indonesia Stock Exchange during the 2022–2024 period (Model 1).

H2a: Working Capital to Total Assets (WCTA) has a significant effect on the Beneish M-Score.

H2b: Retained Earnings to Total Assets (RETA) has a significant effect on the Beneish M-Score.

H2c: Earnings Before Interest and Taxes to Total Assets (EBITA) has a significant effect on the Beneish M-score.

H2d: The market Value of Equity to Total Liabilities (MVTL) has a significant effect on the Beneish M-score.
 H2e: Net Sales to Total Assets (NSTA) has a significant effect on the Beneish M-Score.

Figure 1 presents the conceptual framework that integrates the proposed hypotheses of this study.

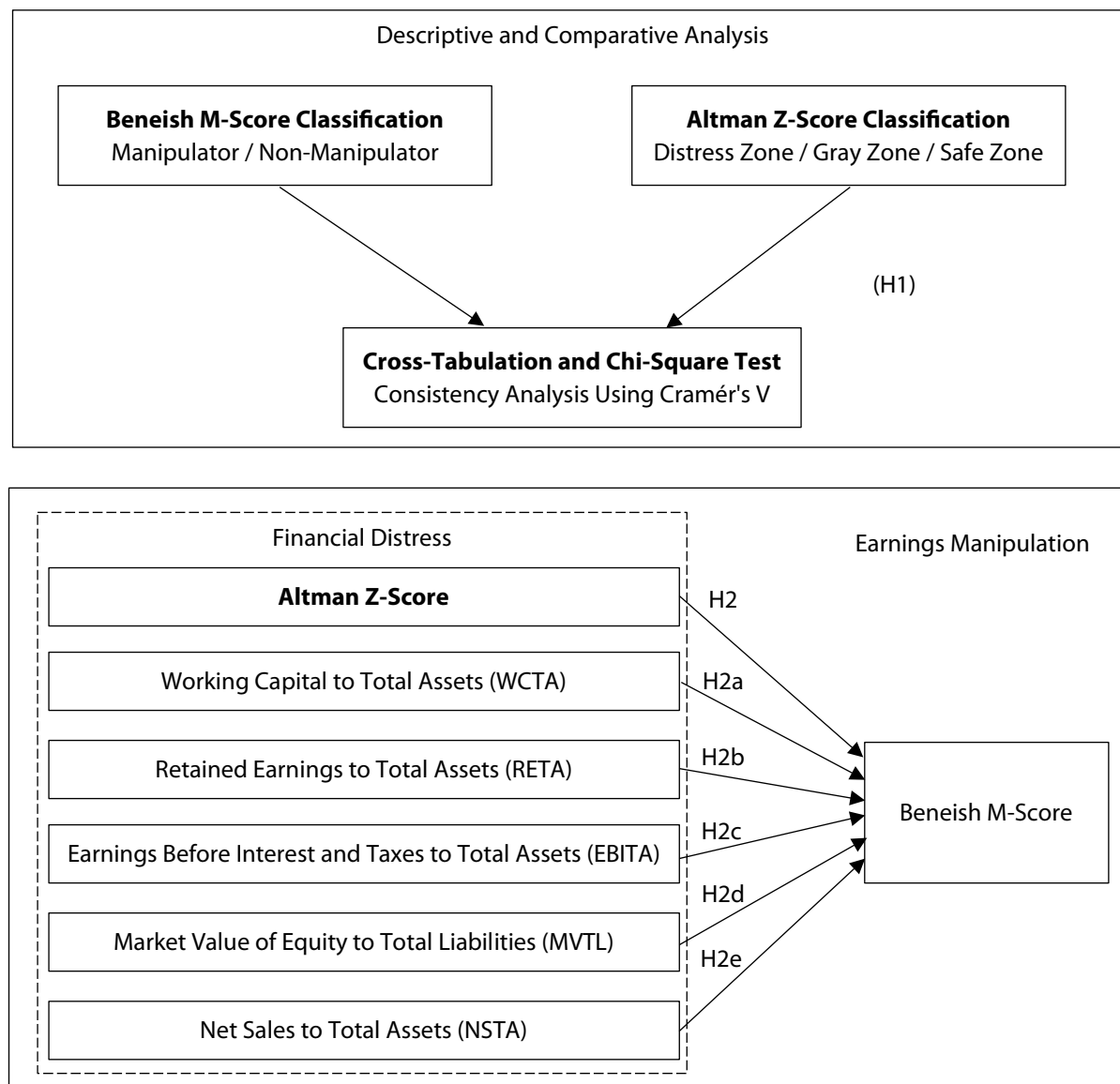


Figure 1. Conceptual Framework

III. Research Method

This study employed a quantitative descriptive-comparative research design. The descriptive component was used to describe the classification results generated by each model, whereas the comparative component examined both the consistency between the two models and the effect of the Altman Z-score on the Beneish M-score. This research framework follows the integrated approach proposed by Kamaluddin et al. (2024) and Valaskova et al. (2021), which has been adapted to the context of the Indonesian energy sector. This study uses secondary data obtained from the annual financial statements of energy sector companies listed on the Indonesia Stock Exchange (IDX), which were retrieved through the Bloomberg Terminal. The data collection period covers 2021 to 2024 because the Beneish M-Score requires financial data from two

consecutive years. Consequently, the effective analysis period was 2022–2024. The study population consists of all companies classified within the Energy Sector under the Indonesia Stock Exchange Industrial Classification (IDX-IC). The sample was selected using purposive sampling based on three criteria (Sugiyono, 2023): (1) the company was listed in the IDX Energy Sector throughout the 2022–2024 period; (2) the company published complete annual reports for all three years; and (3) the company disclosed all financial information required to calculate the eight Beneish M-score ratios and five Altman Z-score ratios. The sample selection process resulted in 61 companies, representing 183 firm-year observations, and forming a balanced panel dataset. All financial ratios were calculated using Microsoft Excel, based on the formulas presented in Sections 2.3 and 2.4. Inferential statistical analyses were performed using Stata version 17.0. Following the panel data framework proposed by Ekananda (2016), the analytical procedure consists of four stages. First, descriptive statistics were generated for all the research variables. Second, the appropriate panel data model was determined by comparing the common, fixed, and random effect models using the Chow and Breusch-Pagan Lagrange Multiplier tests. Third, classical assumption tests were conducted, including the Shapiro-Wilk test for normality, the Variance Inflation Factor (VIF) for multicollinearity, the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity, and the Modified Bhargava et al. Durbin-Watson test together with the Baltagi-Wu LBI test for autocorrelation. In addition, firms were classified according to the threshold values of the Beneish M-Score and Altman Z-scores. Finally, hypothesis testing was performed using the Chi-square test and Cramér's V to test H1, followed by Pooled Ordinary Least Squares (Pooled OLS) regression with robust standard errors to test H2 and H2a through H2e. Two regression models were estimated as follows:

- Model 1:

$$MSCORE_{it} = \alpha + \beta_1 ZSCORE_{it} + \epsilon_{it}$$

- Model 2:

$$MSCORE_{it} = \alpha + \beta_1 WCTA_{it} + \beta_2 RETA_{it} + \beta_3 EBITA_{it} + \beta_4 MVTL_{it} + \beta_5 NSTA_{it} + \epsilon_{it}$$

where MSCORE represents the Beneish M-Score of company *i* in year *t*, ZSCORE represents the aggregate Altman Z-Score, and WCTA, RETA, EBITA, MVTL, and NSTA represent the five financial ratio components of the Altman Z-score. The error term is denoted by ϵ_{it} . Robust standard errors were employed to correct for heteroscedasticity identified during the diagnostic testing stage.

Table 1. Sample Selection (N = 183)

No.	Criteria	Companies	Firm-Year Observations
1	Total companies listed in the IDX Energy Sector	91	-
2	Companies without complete annual reports for the 2022–2024 period	(5)	-
3	Companies with complete annual reports	86	250
4	Companies with incomplete data required for the Beneish M-Score and Altman Z-Score calculations	(25)	(67)
5	Final sample	61	183

IV. Results and Discussion

4.1. Descriptive Statistics

Table 2 presents the descriptive statistics for the 15 variables included in the analysis. The mean Altman Z-Score is 2.9448, which is slightly below the safe zone threshold of 2.99. This finding indicates that, on average, the energy companies in the sample are positioned near the boundary between the grey and safe

zones. The relatively large standard deviation of 4.8103, which exceeds the mean value, indicates substantial heterogeneity among the sample firms, with Z-scores ranging from -15.7798 to 23.8506 . Meanwhile, the mean Beneish M-Score is -1.9079 , which is above the manipulation threshold of -2.22 , suggesting that, on average, the sample firms are classified as potential earnings manipulators.

Table 2. Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
ZSCORE	183	2.9448	4.8103	-15.7798	23.8506
MSCORE		-1.9079	2.4445	-4.9973	20.6482
WCTA (X1)		0.0622	0.3498	-1.4157	0.6961
RETA (X2)		-0.2111	1.5773	-11.4580	0.7493
EBITA (X3)		0.1289	0.1642	-0.1291	0.8120
MVTL (X4)		3.2111	5.4384	0.0219	34.5246
NSTA (X5)		0.8138	0.6148	0.0396	3.2654
DSRI		1.1582	0.9436	0.0578	10.7124
GMI		1.2766	2.7639	-7.1723	36.9505
AQI		1.5864	2.7094	0.0044	23.2977
SGI		1.2886	1.3764	0.2394	14.8906
DEPI		1.0413	0.4257	0.2651	3.5292
SGAI		1.0827	0.4848	0.1145	4.6597
TATA		-0.0452	0.1213	-0.6488	0.8028
LVGI		0.9778	0.2286	0.2910	2.4556

Source: Author's own calculations using Stata 17.0.

Among the components of the Altman Z-Score, RETA has a negative mean value of -0.2111 , indicating that many firms in the sample have accumulated their retained losses. In contrast, EBITA has a positive mean value of 0.1289 , suggesting that many companies maintained positive operating profitability during the commodity boom period. The wide range of MVTL values, from 0.0219 to 34.5246 , reflects the substantial differences in market valuation across the sampled firms. With respect to the Beneish M-Score variables, both the DSRI and GMI have mean values greater than 1, indicating that, on average, the collection period for receivables increased while gross profit margins declined compared with the previous year. According to the Beneish model, these patterns are potential indicators of earnings manipulation.

4.2. Panel Model Selection

Two statistical tests were conducted to determine the most appropriate panel data model for regression analysis. The Chow test was used to compare the Common Effect Model (CEM) with the Fixed Effect Model (FEM), whereas the Breusch-Pagan Lagrange Multiplier (LM) test was employed to compare the Common Effect Model (CEM) with the Random Effect Model (REM). The results are shown in Table 3.

Table 3. Panel Model Selection Tests

Test	Statistic	p-value	Decision
Chow Test (CEM vs. FEM)	$F(60, 121) = 0.76$	0.8753	Fail to reject H_0 ; CEM is preferred
Breusch-Pagan LM Test (CEM vs. REM)	$\chi^2 = 0.00$	1.0000	Fail to reject H_0 ; CEM is preferred
Selected Model	-	-	Common Effect Model (Pooled OLS)

The results of both model selection tests consistently support the Common Effect Model. Because the Common Effect Model was preferred over both the Fixed Effect Model and the Random Effect Model, the

Hausman test was not required. These findings indicate that firm-specific effects are not statistically significant in this sample. Therefore, the Common Effect Model, estimated using Pooled Ordinary Least Squares (Pooled OLS), provides the most appropriate specification for examining the relationship between the Altman Z-score and the Beneish M-score.

4.3. Classical Assumption Tests

The Shapiro-Wilk test produced a test statistic of $W = 0.5086$ with a p-value of 0.0000, indicating that the regression residuals were not normally distributed. However, because the sample consists of 183 observations, the Central Limit Theorem supports proceeding with the analysis, as the sampling distribution of the estimator approaches normality when the sample size is sufficiently large, regardless of the underlying distribution of residuals (Ghozali, 2018). The multicollinearity diagnostics for Model 2 indicate a maximum Variance Inflation Factor (VIF) of 1.92 for EBITA and a mean VIF of 1.61, both of which are well below the commonly accepted threshold of 10. In addition, all tolerance values exceeded 0.10. These results indicate that multicollinearity is not a concern and that the five explanatory variables can be interpreted independently of each other. The Breusch-Pagan/Cook-Weisberg test yielded $\chi^2(1) = 13.01$ with a p-value of 0.0003, indicating heteroscedasticity. To address this issue, White's heteroscedasticity-consistent robust standard errors were employed, a procedure that is widely accepted in modern econometric analysis (White, 1980; Wooldridge, 2010). To assess autocorrelation, the Modified Bhargava et al. The Durbin-Watson statistics were 1.7402 for Model 1 and 1.7854 for Model 2, both of which were close to the benchmark value of 2. Similarly, the Baltagi-Wu LBI statistics were 2.4935 and 2.5236, while the estimated AR(1) coefficients were 0.1299 and 0.1092, respectively. These results indicate that serial correlation is not a significant issue in either of the regression models. Overall, the diagnostic tests support the use of the Common Effect Model estimated by Pooled Ordinary Least Squares (OLS) with robust standard errors.

4.4. Beneish M-Score Classification Results

Table 4 presents the annual classification results based on Beneish M-scores. Across the 2022–2024 observation period, 71 of the 183 firm-year observations (38.80%) were classified as potential manipulators, whereas 112 observations (61.20%) were classified as non-manipulators. The yearly pattern revealed substantial variations. The proportion of potential manipulators increased from 40.98% in 2022 to 52.46% in 2023, before declining sharply to 22.95% in 2024.

Table 4. Beneish M-Score Classification by Year

Classification	2022	2023	2024	Total (2022–2024)
Manipulator ($M > -2.22$)	25 (40.98%)	32 (52.46%)	14 (22.95%)	71 (38.80%)
Non-Manipulator ($M \leq -2.22$)	36 (59.02%)	29 (47.54%)	47 (77.05%)	112 (61.20%)
Total	61	61	61	183

The sharp increase in the proportion of potential manipulators in 2023 can be attributed to the commodity supercycle. Exceptional growth in corporate revenues increased both the Sales Growth Index (SGI) and the Gross Margin Index (GMI), causing these ratios to deviate substantially from their normal levels. According to the Beneish M-Score model, these unusual movements are potential indicators of earnings manipulation. In 2024, the substantial decline in coal prices, which approached 62%, contributed to the normalization of most financial ratios, resulting in a lower proportion of firms being classified as potential manipulators. These findings suggest that the Beneish M-score classification in highly volatile industries should be interpreted within the broader macroeconomic environment. Furthermore, three companies were consistently classified as potential manipulators throughout 2022–2024, indicating persistent reporting patterns that warrant further attention. The overall detection rate of 38.80% is broadly consistent with the

findings of Kamaluddin et al. (2024), who reported a detection rate of 39.53% among Malaysian listed companies, and is slightly higher than the 36.00% reported by Islamiyati and Purnomo (2021) for Indonesian manufacturing firms. Overall, the Beneish M-Score effectively distinguishes between potential manipulators and non-manipulators within the Indonesian energy sector, with the observed classification pattern closely reflecting fluctuations in commodity prices.

4.5. Altman Z-Score Classification Results

Table 5 presents the annual classification results based on Altman's Z-score. Across the 2022–2024 observation period, 78 firm-year observations (42.62%) were classified in the distress zone, 38 observations (20.77%) were classified in the gray zone, and 67 observations (36.61%) were classified in the safe zone. Unlike the Beneish M-score classifications, which fluctuated considerably across the observation period, the Altman Z-score classifications remained relatively stable. The distributions for 2022 and 2024 were nearly identical, with only slight variations observed in 2023.

Table 5. Altman Z-Score Classification by Year

Zone	2022	2023	2024	Total (2022–2024)
Distress ($Z < 1.81$)	25 (40.98%)	28 (45.90%)	25 (40.98%)	78 (42.62%)
Gray ($1.81 \leq Z \leq 2.99$)	13 (21.31%)	12 (19.67%)	13 (21.31%)	38 (20.77%)
Safe ($Z > 2.99$)	23 (37.70%)	21 (34.43%)	23 (37.70%)	67 (36.61%)
Total	61	61	61	183

The relative stability of the Altman Z-score classification reflects the characteristics of the model itself. The financial ratios underlying the Z-score, including retained earnings, capital structure, and operating profitability, generally change gradually rather than responding immediately to fluctuations in commodity prices. Among the sampled companies, 22 firms remained in the distress zone throughout the entire observation period, indicating persistent structural financial weaknesses, rather than temporary cyclical difficulties. From the perspective of the Fraud Triangle Theory (Cressey, 1953), the large proportion of firms classified in the distress zone indicates that financial pressure is prevalent within the Indonesian energy sector. As demonstrated by Rosner (2003), firms experiencing financial distress have stronger incentives to manipulate earnings than financially healthy firms.

4.6. Cross-Tabulation, Chi-Square Test, and Cramér's V

The first hypothesis (H1) was tested using cross-tabulation between the Altman Z-score and Beneish M-score classifications. The association between the two classification models was evaluated using the Pearson Chi-square test and Cramér's V. The results are presented in Tables 6 and 7, respectively.

Table 6. Cross-Tabulation of Altman Z-Score and Beneish M-Score Classifications

Altman Z-Score Zone	Non-Manipulator ($M \leq -2.22$)	Manipulator ($M > -2.22$)	Total
Distress Zone ($Z < 1.81$)	49 (62.82%)	29 (37.18%)	78
Gray Zone ($1.81 \leq Z \leq 2.99$)	31 (81.58%)	7 (18.42%)	38
Safe Zone ($Z > 2.99$)	32 (47.76%)	35 (52.24%)	67
Total	112 (61.20%)	71 (38.80%)	183

Table 7. Chi-Square Test and Cramér's V Results

Zone	N	Actual Manipulators	Expected Manipulators	Difference	Manipulator Percentage	Interpretation
Distress	78	29	30.3	-1.3	37.18%	Close to expected
Gray	38	7	14.7	-7.7	18.42%	Substantially below the expected value
Safe	67	35	26.0	+9.0	52.24%	Substantially above the expected value
Total	183	71	71.0	-	38.80%	Overall average
Pearson χ^2	11.83	df = 2	p = 0.003	-	-	Significant
Cramér's V	0.2542	-	-	-	-	Moderately strong

The Pearson Chi-square test produced a value of $\chi^2 = 11.83$ with a p-value of 0.003, indicating a statistically significant association between the Altman Z-Score and Beneish M-Score classifications. In addition, Cramér's V coefficient of 0.2542 indicates a moderate level of association according to conventional effect size guidelines. Therefore, H1 is supported in this study. This finding is consistent with previous studies by Kamaluddin et al. (2024) and Valaskova et al. (2021), who also reported that the two models measure different financial constructs but produce systematically related classification outcomes. A more noteworthy finding emerged from the distribution of the classification results. Contrary to the prediction of the Fraud Triangle Theory, the proportion of potential manipulators does not consistently increase with the level of financial distress. Instead, the distribution followed a U-shaped pattern. The proportion of potential manipulators was 37.18% in the distress zone, declined to 18.42% in the grey zone, and increased substantially to 52.24% in the safe zone. Consequently, the highest proportion of potential manipulators is observed among firms classified as being in the safe zone. This pattern suggests that different mechanisms may operate at the opposite ends of the financial condition spectrum. The pressure element of the Fraud Triangle Theory remains relevant among firms in the distress zone. Companies experiencing financial difficulties may manipulate earnings to avoid bankruptcy signals, comply with debt covenants, or maintain the confidence of creditors and investors. Rosner (2003) documented similar behavior among financially distressed firms approaching bankruptcy in the U.S..

In contrast, firms in the safe zone appear to be motivated by different mechanisms. Consistent with Agency Theory (Jensen & Meckling, 1976), managers of financially healthy firms may have incentives to maintain favorable growth expectations through income smoothing or aggressive accounting practices. During the 2022–2024 commodity supercycle, many energy companies experienced exceptional revenue growth. Combined with managerial discretion in financial reporting, these unusually strong increases produce extreme values for the Sales Growth Index (SGI) and Gross Margin Index (GMI), causing the Beneish M-score to identify potential manipulation even when the underlying business performance is primarily driven by favorable market conditions. Firms in the gray zone, positioned between financial distress and financial strength, appear to face weaker incentives for either defensive earnings manipulation or aggressive income reporting, making conservative financial reporting the most reasonable strategy. Among the 183 firm-year observations, 29 (15.85%) were simultaneously classified as financially distressed and potential manipulators, whereas 32 (17.49%) were simultaneously classified as financially healthy and non-manipulators. When the gray zone was treated as a transitional category, the overall concordance rate between the two models was approximately 33.33%. These findings suggest that the Beneish M-Score-and Altman Z-scores should be viewed as complementary diagnostic tools rather than interchangeable measures. This conclusion is

consistent with the multiple model approach recommended by Kamaluddin et al. (2024) and Valaskova et al. (2021).

4.7. Regression Analysis (H2 and H2a–H2e)

Table 8 presents the regression results for Model 1, which uses the aggregate Altman Z-score as the independent variable, and Model 2, which includes the five individual components of the Altman Z-score. Both models were estimated using Pooled Ordinary Least Squares (POLS) with robust standard errors.

Table 8. Pooled OLS Regression Results (Robust Standard Errors)

Variable	Model 1: Coefficient	Model 1: p-value	Model 2: Coefficient	Model 2: p-value
Constant	-1.924	0.000	-1.573	0.000
ZSCORE	0.005	0.851	-	-
WCTA	-	-	0.511	0.559
RETA	-	-	0.070	0.409
EBITA	-	-	-1.145	0.397
MVTL	-	-	0.001	0.958
NSTA	-	-	-0.254	0.347
N	183	-	183	-
R ²	0.0001	-	0.0208	-
F-statistic	F(1,181) = 0.04	0.8514	F(5,177) = 1.24	0.2921

The results indicate that neither of the regression models produced a statistically significant relationship. In Model 1, the coefficient of the aggregate Altman Z-Score is 0.005, with a p-value of 0.851, indicating that financial distress, as measured by the aggregate Z-Score, does not significantly influence the Beneish M-Score. Likewise, in Model 2, none of the five financial ratio components of the Altman Z-Score are statistically significant, with p-values ranging from 0.347 to 0.958. In addition, the overall regression model is not statistically significant, as indicated by an F-statistic of 1.24, with a p-value of 0.2921. The coefficients of determination (R²) are 0.0001 for Model 1 and 0.0208 for Model 2, indicating that the linear regression models explain virtually none of the variation in the M-score. Therefore, H2, H2a, H2b, H2c, H2d, and H2e were not supported. These findings differ substantially from those reported by Kamaluddin et al. (2024), who found that the aggregate Altman Z-score and four of its five financial ratio components had statistically significant effects on the Beneish M-score using a sample of 3,242 firm-year observations from Malaysia (R² = 0.008; F = 27.548; p < 0.01). Several factors may explain this difference. First, the research contexts differ considerably. Kamaluddin et al. (2024) analyzed companies from multiple industries in Malaysia from 2014 to 2018, whereas the present study focuses exclusively on Indonesian energy companies during the 2022–2024 commodity supercycle.

The energy sector has distinct accounting characteristics that may weaken the conventional relationship between financial distress indicators and earnings manipulation. Second, the relationship between financial distress and earnings manipulation in this study appears to be nonlinear rather than linear. As demonstrated by the chi-square analysis presented in Section 4.6, the proportion of potential manipulators follows a U-shaped pattern, with relatively high proportions observed among both financially distressed and financially healthy firms. Because linear regression assumes a monotonic relationship between independent and dependent variables, it cannot adequately capture this U-shaped pattern. Consequently, the regression coefficients approach zero, not because no relationship exists, but because the underlying relationship cannot be represented by a simple linear model. Therefore, insignificant regression results are more likely to reflect model misspecification than the absence of a meaningful relationship. Third, differences in sample sizes may also contribute to contrasting findings. The present study analyzes 183 firm-year observations, whereas Kamaluddin et al. (2024) examined 3,242 firm-year observations. According to Hair et al. (2019), detecting very

small effect sizes, particularly when R^2 is below 0.02, generally requires substantially larger samples to achieve adequate statistical power. Taken together, the insignificant regression results and the statistically significant chi-square test provide strong evidence that the relationship between financial distress and earnings manipulation among Indonesian energy companies is non-linear. These findings do not invalidate the analytical framework proposed by Kamaluddin et al. (2024). Instead, they suggest that while the framework remains appropriate for broad cross-industry analyses, its application to highly volatile and sector-specific contexts requires methodological adjustments that account for potential nonlinear relationships.

Table 9. Summary of Hypothesis Testing

Hypothesis	Statement	Result
H1	There was a significant association between the Altman Z-score and Beneish M-score classifications.	Supported
H2	The aggregate Altman Z-score has a significant effect on the Beneish M-score.	Not Supported
H2a	Working Capital to Total Assets (WCTA) has a significant effect on the Beneish M-Score.	Not Supported
H2b	Retained Earnings to Total Assets (RETA) has a significant effect on the Beneish M-Score.	Not Supported
H2c	Earnings Before Interest and Taxes to Total Assets (EBITA) has a significant effect on the Beneish M-Score.	Not Supported
H2d	Market Value of Equity to Total Liabilities (MVTL) has a significant effect on the Beneish M-Score.	Not Supported
H2e	Net Sales to Total Assets (NSTA) has a significant effect on the Beneish M-Score.	Not Supported

V. Conclusion

This study examined the relationship between the Beneish M-Score and the Altman Z-Score in detecting fraudulent financial reporting among 61 energy companies listed on the Indonesia Stock Exchange (IDX) during the 2022 to 2024 period. The Beneish M-Score classified 71 of the 183 firm-year observations (38.80%) as potential manipulators, with annual detection rates ranging from 22.95% in 2024 to 52.46% in 2023. This temporal pattern is consistent with fluctuations in global commodity prices, suggesting that the Beneish M-score classifications in the energy sector should be interpreted within the broader macroeconomic environment. Meanwhile, the Altman Z-Score classified 78 observations (42.62%) in the distress zone, 38 observations (20.77%) in the grey zone, and 67 observations (36.61%) in the safe zone, with 22 companies remaining in the distress zone throughout the three-year observation period. The relative stability of the Altman Z-score, compared with the greater variability of the Beneish M-score, reflects the different financial dimensions captured by the two models. The first major finding was that the Pearson Chi-square test identified a statistically significant association between the two classification models ($\chi^2 = 11.83$, $p = 0.003$; Cramér's $V = 0.254$). More importantly, this relationship follows a U-shaped pattern rather than the monotonic pattern predicted by the Fraud Triangle Theory. The highest proportion of potential manipulators was observed in the safe zone (52.24%), followed by the distress zone (37.18%), and the gray zone recorded the lowest proportion (18.42%). This finding suggests that two distinct mechanisms influence earnings manipulation in the Indonesian energy sector. Financially distressed firms may manipulate earnings because of financial pressure, whereas financially healthy firms may engage in aggressive accounting practices to maintain favorable growth expectations and satisfy their market performance objectives. The second major finding is that neither regression model identified a statistically significant linear relationship between the Altman Z-score, including its individual components, and the Beneish M-score. Rather than disproving the theoretical relationship

between financial distress and earnings manipulation, these results indicate that a linear regression model cannot capture the non-linear U-shaped relationship identified by the chi-square analysis.

Consequently, the findings extend the Fraud Triangle Theory by demonstrating that financial pressure alone cannot fully explain earnings manipulation in volatile industries. Agency Theory provides an additional explanation by suggesting that managers of financially healthy firms may also have incentives to manipulate financial reporting to sustain favorable growth narratives and preserve investor confidence. From a practical perspective, the findings suggest that investors, creditors, regulators, and other stakeholders should not limit fraud-risk assessments to financially distressed companies. Energy companies classified within the safe zone may also present a substantial risk of manipulating earnings, particularly during periods of exceptional commodity price growth. Therefore, combining the Beneish M-Score and Altman Z-Score as complementary screening tools is likely to provide a more comprehensive assessment of fraudulent financial reporting risk than relying on either model individually. This study had several limitations. First, the analysis is limited to the Indonesian energy sector and a relatively small sample of 183 firm-year observations, which may restrict the generalizability of the findings and reduce the statistical power to detect small effects. Second, the 2022–2024 observation period was characterized by post-pandemic economic recovery and significant geopolitical disruptions, which may limit the applicability of the findings to more stable economic conditions. Third, the regression analysis employed only linear model specifications, whereas the results indicate that the relationship between financial distress and earnings manipulation is nonlinear. Future research should examine this relationship using non-linear approaches, such as quadratic regression, logistic regression, or non-parametric methods. Future studies may also expand the analysis to other industrial sectors and incorporate additional control variables, including firm size, leverage, corporate governance characteristics, and auditor quality, to improve the explanatory power and robustness of the empirical models.

References

- Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *The Journal of Finance*, 23.
- Association of Certified Fraud Examiners (ACFE). (2024). *The nations occupational fraud 2024: A report to the nations*. ACFE.
- Baltagi, B. H. (2021). *Econometric analysis of panel data*. Springer.
- Beneish, M. D. (1999). Detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Chenkiani, P., & Prasetyo, A. (2023). *Fraud Dan Monitoring Dalam Perspektif Teori Keagenan*. Agustus, 12(2). <https://doi.org/10.46806/ja.v12i1.1016>
- Christian, N., Istiqomah, N., Simdy, V., & Dharmawan, C. (2024). Analisis kasus PT Bumi Resources Tbk dengan teknik cash flow shenanigans. *Angel Innovative: Journal of Social Science Research*, 4(4), 3768–3776.
- Cressey, D. R. (1953). *Other people's money: A study of the social psychology of embezzlement*. Free Press.
- Dinasmara, C. K., & Adiwibowo, A. S. (2020). Deteksi kecurangan laporan keuangan menggunakan Beneish M-Score dan prediksi kebangkrutan menggunakan Altman Z-Score (Studi empiris pada perusahaan yang termasuk dalam indeks LQ-45 tahun 2016–2018). *Diponegoro Journal of Accounting*, 9(3), 1–15.
- Ekananda, M. (2016). *Analisis ekonometrika data panel (2nd ed.)*. Mitra Wacana Media.
- Ghozali, I. (2018). *Aplikasi analisis multivariate dengan program IBM SPSS 25 (9th ed.)*. Badan Penerbit Universitas Diponegoro.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics (5th ed.)*. Douglas Reiner.
- Hair, J. F. ., Black, W. C. ., Babin, B. J. ., & Anderson, R. E. . (2019). *Multivariate data analysis by Joseph F. Hair, Jr. (University of South Alabama), William C. Black (Louisiana State University) and Barry J. Babin (Louisiana Tech University), Rolph E. Anderson (Drexel University)*. Cengage.
- Handoko, B. L., Wiyardi, E. S., & Handoko, J. S. (2022). The application of Beneish M-Score method in detecting fraudulent manipulation on financial statements (Case study of Indonesian government-state owned

- enterprise (SOE) registered on Indonesian Stock Exchange 2016–2020). ACM International Conference Proceeding Series, 457–466. <https://doi.org/10.1145/3556089.3556106>
- Herianti, E., Suryani, A., & Marundha, A. (2023). Audit kecurangan laporan keuangan (1st ed.). Eureka Media Aksara.
- Islamiyati, D., & Purnomo, D. E. (2021). Komparasi akurasi model M-Score dan Z-Score dalam mendeteksi fraudulent financial reporting. *Jurnal Neraca*, 1, 1–18. <https://doi.org/10.48144/neraca.v17i1.591>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics*, 3: 305–360.
- Jiang, J. (2023). A synthesized distribution model: Asymmetric information, agency problems, and intertemporal optimization. *Corporate Governance and Organizational Behavior Review*, 7(4), 152–160. <https://doi.org/10.22495/cgobrv7i4p13>
- Kamaluddin, A., Boni, L., Sutainim, N. A., & Mohammed, N. F. (2024). The relationship between Z-Score and Beneish M-Score: Evidence from Malaysian public listed companies. *Journal of Business and Social Development*, 12(2), 16–35.
- MacCarthy, J. (2017). Using Altman Z-Score and Beneish M-Score models to detect financial fraud and corporate failure: A case study of the Enron Corporation. *International Journal of Finance and Accounting*, 6(6), 159–166. <https://doi.org/10.5923/j.ijfa.20170606.01>
- Marini, S. S., Saputri, R. I., Muthi'ah Ashari, F., Kusumaningtias, R., & Kusumaningsih, A. (2025). Analisis penerapan prinsip tata kelola korporat pada kasus manipulasi laporan keuangan: Studi kasus PT Cakra Mineral Tbk dan PT Bumi Resources Tbk. *Inovasi dan Kreativitas dalam Ekonomi*, 8(5), 134–141.
- Miharsi, D. Gamayuni, R. R., & Dharma, F. (2024). Analysis of the utilization of Altman Z-Score, Beneish M-Score, and F-Score models in detecting fraudulent financial reporting: A literature review. *Journal of Management, Accounting, General Finance and International Economic Issues (MARGINAL)*, 3(2). <https://ojs.transpublika.com/index.php/MARGINAL/>
- Pertiwi, J. C., Oktavia, R., & Amelia, Y. (2023). Analisis perbandingan metode pendeteksian kecurangan keuangan menggunakan Altman Z-Score, Beneish M-Score, dan Springate. *Jurnal Ilmiah Akuntansi dan Keuangan*, 5(6).
- Putra, Y. P. (2021). Perbandingan metode Altman Z-Score, Beneish M-Score – Data Mining dan Springate dalam mendeteksi fraudulent financial reporting (Studi empiris perusahaan manufaktur tahun 2014–2018). *EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi dan Bisnis*, 9(1), 81–94. <https://doi.org/10.37676/ekombis.v9i1.1222>
- Rosner, R. L. (2003). Earnings manipulation by failing firms. *Contemporary Accounting Research*, 20(2), 361–408. <https://doi.org/10.1506/8EVN-9KRB-3AE4-EE81>
- Rusci, V. A., Santosa, S., and Fitriana, V. E. (2021). Financial distress and earnings management in Indonesia: The role of independent commissioners. *Jurnal Ilmiah Akuntansi Fakultas Ekonomi*, 7(1), 89–104. <https://doi.org/10.34204/jiafe.v7i1.3153>
- Saputri, N. L., & Irene, K. F. (2024). Mengungkap tanda-tanda skandal korupsi 271 T PT Timah Tbk, analisis earning opacity. *Jurnal Penelitian Ilmu-Ilmu Sosial*, 2(5). <https://doi.org/10.5281/zenodo.14580007>
- Sihombing, P. R. (2022). Aplikasi STATA Untuk Statistisi Pemula. <https://www.researchgate.net/publication/358460661>
- Sugiyono. (2023). Metode Penelitian Kuantitatif, Kualitatif, Dan R&D. Alfabeta Bandung. www.cvalfabeta.com
- Suheni, V., & Arif, M. F. (2020). Mendeteksi financial statement fraud dengan menggunakan model Beneish M-Score (Studi pada perusahaan sektor manufaktur yang terdaftar di Bursa Efek Indonesia). *Jurnal Akuntansi & Ekonomi FE UN PGRI Kediri*, 5(2).
- Tanusdjaja, H., & Kurniawan, F. M. (2018). Analisis komparasi metode Altman Z-Score – Financial Ratio dan metode Beneish M-Score Model – Data Mining dalam mendeteksi fraudulent financial reporting. *Jurnal Muara Ilmu Ekonomi dan Bisnis*, 2(1).

- Valášková, K., Androniceanu, A. M., Zvariková, K., & Oláh, J. (2021). Bonds between earnings management and corporate financial stability in the context of enterprises' competitive ability. *Journal of Competitiveness*, 13(4), 167–184. <https://doi.org/10.7441/JOC.2021.04.10>
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838. <https://doi.org/10.2307/1912934>
- Wolfe, D. T., & Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *The CPA Journal*, 74(12), 38–42.
- Wooldridge, J. M. (2010). *Econometric analysis of cross-section and panel data* (2nd ed.). MIT Press.