Beneish M-Score Reliability as a Tool For Detecting Financial Statements Fraud

Ria Septiani[1], Siti Musyarofah[2], Rita Yuliana[3]
[1]riaseptiani973@gmail.com, [2]sitimusyarofah@trunojoyo.ac.id, [3]rita.yuliana@trunojoyo.ac.id
[1]Universitas Trunojoyo Madura

Abstract: The rise of fraud has made a reliable and precise fraud detection model based on financial reports critical to developing. This study uses discriminant analysis to test the Beneish M-Score's ability in detecting fraud in the presentation of financial statements with a sample of 114 financial statements of banking companies for 2016-2018. The study results using the discriminant analysis method found that the beneish m-score was able to detect fraud by 89.5%. Meanwhile, the Beneish DSRI, GMI, AQI, DPI, and TATA ratios prove significant in grouping companies into manipulators and non-manipulators. This research concludes that the Beneish M-Score model is accurate in detecting fraud in its financial statements. The Beneish M-Score ratios contributing to grouping financial reports into the companies' manipulator and non-manipulator companies are DSRI, GMI, AQI, DPI, and TATA. DSRI is the most dominant ratio in grouping these companies.

Keywords: Fraud; Benesih M-score; Discriminant analysis

Introduction

One of the critical issues in current accounting research is the behavior of company managers who tend to submit financial reports that are favorable to their side (Beneish 2001). Many acts of fraud, manipulation of financial statements, and unethical actions by companies have shocked the global economy of late. Some of the fraud scandals were carried out by large companies such as Enron, Xerox, Worldcom, and phenomenal fraud cases by financial companies Lehman Brothers and AIG. These fraud cases continue to increase in number from year to year (Kasim and Higson 2012).

The banking industry in Indonesia was also surprised by several findings of fraud by companies. OJK (Otoritas Jasa Keuangan), as the institution authorized to supervise banking companies in Indonesia reported that there were 119 cases of fraud in the banking sector during 2016-2018, with details of 26 cases in 2016, 57 cases in 2017 and 2018, 36 a case involving banking. During the last three years, the types of issues included 55% of credit cases, 21% of fraudulent records, 15% of embezzlement of funds, 5% of fund transfers, and 4% of asset procurement. From the OJK survey results, 90% of repair fraud occurred because it involved people within the companies.

The Bank Bukopin case is one of Indonesia's major fraud cases, where the company allegedly revamped its 2015, 2016, and 2017 financial statements by inflating a profit of Rp. 1.08 trillion. The details of the fraud committed by Bank Bukopin include the alleged net profit, which should have been reported at Rp. 183.56 billion turned out to be inflated to Rp. 1.08 trillion, the provision income from credit cards should be Rp. 317.88 billion to Rp. 1.06 trillion, the allowance for impairment losses on assets should be Rp. 649.05 billion to Rp. 797.65 billion, so that there was an increase in company expenses by Rp. 148.6 billion. For fraud committed by Bank Bukopin, it was administratively punished and fined.

Several other fraud cases were also revealed to be committed by banking companies, such as the fictitious credit case carried out by Bank BJB (Regional

Development Bank of West Java and Banten), which was revealed in 2017 as a case of notional credit worth 548 billion (financialbisnis.com).

Through research, the world of education can help OJK's task in early detection of fraud based on financial reports that have been released by the company. From the company's financial statements, an analysis can be carried out to get a score of whether the company can be classified as a manipulator company that tends to commit fraudulent financial statements. The score, known as the Beneish M-Score, was developed by Professor M. Daniel Beneish. This model involves several financial ratios to get a specific score to identify the possibility of fraud in preparing the company's financial statements. These ratios include receivable days sales index, gross margin index, asset quality index, grow sales index, depreciation index, general and administration sales index, leverage index, total accrual to total assets. Based on the score obtained, a company can be categorized as a manipulator and non-manipulator group. The Beneish M-Score itself is a probabilistic model (Beneish 1999).

Several studies were carried out using the Beneish M-Score approach to detect fraudulent financial statements at companies from various countries, which showed the Beneish M-Score's effectiveness in detecting fraud (Kartika and Irianto 2010;Omar et al 2014; Tarjo and Herawati 2014; Mahama 2015). However, several other studies have found that the Beneish M-Score is not effective in detecting financial report fraud (Gyarteng 2014; Mehta anad Bhavani 2017; Bhavani and Amponsah 2017; Santoso and Ginting 2019).

Based on the rampant phenomena of fraud cases that have recently involved banking companies in Indonesia. And there is still a gap in the results of testing the effectiveness of the M-Score model in detecting financial report fraud. This study aims to test the Beneish M-Score model to analyze financial report fraud in Indonesia's banking sector.

Literature Review

Fraudulent of Financial Statements

Fraud is the misuse of corporate power or assets deliberately to enrich oneself (ACFE, Fraud Resources, 2018). As a result of frau actions, some parties get personal benefits, and on the other hand, some parties are harmed, be it companies or other parties. The fraud scheme that is carried out or known as the Fraud Tree is corruption, misuse of assets, and fraudulent financial statements (Arens et al. 2012).

Fraudulent financial statements are misstatements or deliberate omissions of amounts or disclosures in financial reports so that users of financial statements have an incorrect perception of the company's financial statements (Albrecht and Zimbelman 2011). Fraudulent financial statements include manipulation, forgery, changes in accounting records, or supporting documents that are the source of data for presenting financial statements. A fraudulent financial statement scheme is in the form of giving net income that is too high (earning overstatement) or too low (earning understatement) by manipulating data on financial statement items and disclosures (ACFE 2018). Earning overstatement is more often done because, with the earning overstatement scheme, the recognition of revenue or profit higher than it should be will increase the company's profitability. Some of the standard practices used to increase or decrease company assets or profits are: (Kartika and Irianto 2010).

- Fraud in the valuation of several company assets such as receivables and inventories, manipulation can also be done in the form of recognizing higher purchase prices for fixed assets and capitalization of stocks that are not true.
- Record fictitious sales transactions, which result in an overstatement of assets and income.
- Undertake understatement of obligations and expenses such as deliberately not recording or hiding transactions related to costs and liabilities.

- Abusing different accounting periods to recognize revenue faster than they should or delay the recognition of expenses.
- Intentionally did not provide accurate information to mislead users of financial statements.

Beneish M-Score

The Beneish M-Score was designed by Professor Messod Beneish, is a financial statement analysis tool to predict fraud or fraud on the company's financial statements. The Beneish M-Score is formulated with eight ratios to identify financial report fraud occurrence or find out that the financial statements have been manipulated (Beneish et al. 2013). The formula for the Beneish M-Score is as follows:

$$\mathrm{M} = (-4.840) + (0{,}920\mathrm{XDSRI}) + (0{,}528\mathrm{XGM}) + (0{,}0404\mathrm{XAQI}) + (0{,}892\mathrm{XSGI}) + (0{,}115\mathrm{XDEPI}) - (0{,}172\mathrm{XSGAI}) + (4{,}679\mathrm{XTATA}) - (0{,}327\mathrm{XLVGI}) \quad (1)$$

The Beneish M-Score formula consists of 8 ratios:

1. *Days Sales in Receivables Index* (DSRI)

DSRI compares trade receivables to sales generated by the company in one year (t) and the previous year (t-1). If DSRI> 1, then this indicates an increase in the number of trade receivables owned (Beneish). This condition means an earning overstatement (Beneish 1999). The formula for DSRI is:

$$DSRI = \frac{\dfrac{Accounts\ receivable_t}{Sales_t}}{\dfrac{Accounts\ receivable_{t-1}}{Sales_{t-1}}} \quad (2)$$

2. *Gross Margin Index (GMI)*

GMI is a ratio that measures the level of the company's profitability, where this ratio represents the company's prospects. If GMI> 1 indicates a decrease in its gross profit, which means a decline in the company's prospects. This condition indicates an earning overstatement (Beneish 1999). The Formula for GMI is:

$$GMI = \frac{\dfrac{(Sales_{t-1} - Cost\ of\ Goods\ Sold_{t-1})}{Sales_{t-1}}}{\dfrac{(Sales_t - Cost\ of\ Goods\ Sold_t)}{Sales_t}} \quad (3)$$

3. *Asset Quality Index (AQI)*

AQI shows the quality of the company's non-current assets that are likely to benefit the company in the future. If AQI> 1, this indicates a decline in asset quality. Thus, if there is an increase in the number of non-current assets that can provide benefits in the future and increase the number of deferred expenses, this condition indicates an earning overstatement (Beneish 1999). The formula for AQI is:

$$AQI = \frac{1 - \dfrac{(Current\ Asset_t + Fixed\ Asset_t)}{Total\ Asset_t}}{1 - \dfrac{(Current\ Asset_{t-1} + Fixed\ Asset_{t-1})}{Total\ Asset_{t-1}}} \quad (4)$$

4. *Sales Growth Index (SGI)*

SGI is the ratio between sales in year t and sales in year t-1. If SGI> 1 indicates an increase in sales, a decrease in this ratio suggests a decline in sales. If SGI> 1, then this means an earning overstatement (Beneish 1999). The formula for SGI is:

$$SGI = \frac{Sales_t}{Sales_{t-1}} \tag{5}$$

5. *Depreciation Index (DEPI)*

Depreciation Index ratio compares the depreciation expense to fixed assets before depreciation in one year (t) and the previous year (t-1). If DEPI> 1 indicates a decrease in the depreciation expense for fixed assets, whereas DEPI <1 shows an increase in fixed assets' depreciation rate. If DEPI> 1, then this means an earning overstatement (Beneish 1999). The formula for DEPI is:

$$DEPI = \frac{\frac{Depreciation_{t-1}}{(Fixed\ asset_{t-1} + Depreciation_{t-1})}}{\frac{Depreciation_t}{(Fixed\ asset_t + Depreciation_t)}} \tag{6}$$

6. *Sales Generation and Administrative Expenses Index (SGAI)*

SGAI ratio compares selling, general, and administrative expenses to sales in one year (t) and the previous year (t-1). If SGAI> 1, then this indicates that there is an increase in company operating expenses - administrative, general and sales expenses or a decrease in sales. And vice versa, if SGAI <1, then this indicates a reduction in a company operating costs or an increase in sales. If SGAI <1, then this means an earning overstatement (Beneish 1999) : (7)

$$SGAI = \frac{\frac{\text{selling, general, and administrative expenses}_t}{Sales_t}}{\frac{\text{selling, general, and administrative expenses}_{t-1}}{Sales_{t-1}}}$$

7. *Total Accruals to Total Assets (TATA)*

Accrued earnings relate to the increase in recognition of company earnings through additional recognition of revenue. High accrual income indicates that the amount of cash on income generated is low. A high TATA ratio indicates the potential condition of the company for earning overstatement through an increase in accrual transactions in revenue recognition (Beneish 1999). TATA can be calculated with the following formula:

$$TATA = \frac{(Gross\ Profit_t - Operating\ cash\ flows_t)}{Total\ Asse_t} \tag{8}$$

8. *Leverage Index (LVGI)*

LVGI is useful for measuring the level of debt a company has against its total assets from year to year. If LVGI> 1, then this indicates an increase in the debt composition of all assets owned by the company, while a decrease in this ratio suggests a reduction in the company's amount of debt. If LVGI> 1, then this indicates the potential condition of the company for earning overstatement to fulfill its obligations (Beneish 1999). The formula for LVGI is:

$$LVGI = \frac{\frac{(Current\ Liabilities_t - \ Longterm\ obligation_t)}{Total\ Asset_t}}{\frac{(Current\ Liabilities_{t-1} - \ Longterm\ obligation_{t-1})}{Total\ Asset_{t-1}}} \tag{9}$$

Research Objectives.
This study classify companies into fraudulent and non-fraudulent groups based on the beneish m-score analysis. Further analyzing each of the eight beneish m-score ratios in the grouping is the ratios that differentiate whether the company is classified as fraudulent or non-fraudulent.

Methods
Fraud is the misuse of corporate power or assets deliberately to enrich oneself (ACFE, 2016). This research is quantitative research with a descriptive approach; the sample used is banking companies listed on the IDX for the period 2016-2018; the sample was taken by purposive sampling, with the following conditions:

- Banking companies listed on the IDX and published financial reports for the period 31 December 2016-2018.
- Issuing financial reports in Rupiah denominations.
- Financial reports have data relating to the variables required in the study

The data used is secondary data obtained through the official IDX website www.idx.go.id. Based on the specified criteria, the research sample was 38 companies with three years, so that the number of samples in this study was 114 companies. A summary of the acquisition of research samples is presented in table 1 as follows:

Table 1. Summary of Sample Selection

Criteria	Number of Data
1. All ofBanking Companies Listed on IDX	43
2. Banking companies that do not publish financial reports for the 2016-2018 period	3
3. Incomplete data	2
Total companies according to the criteria	38
Total financial reports in the study period (2016-2018)	114

Findings
Descriptive Research Samples
Researchers have conducted a beneish m-score-based analysis of several 114 financial reports (2016-2018) from 38 companies that were the research samples. Where a total of 28 samples were classified as fraudulent, and 86 samples were classified as non-fraudulent. Table 2 presents descriptions of each of the benefit ratios of the entire study sample. Meanwhile, table 3 and table 4 present a descriptive description of each fraudulent and non-fraudulent group.

Table 2. Descriptive Statistics of Beneish M-Score Ratio All Research Samples

Deskriptif	DSRI	GMI	AQI	SGI	DEPI	TATA	LVGI
Median	1.0476132	0.95777408	0.99956172	1.04556057	0.950902	0.002998	0.985881
Std. Dev	53.39773	1.57063155	0.64592673	2.5624726	1.428614	0.248179	9.731474
Minimum	-1.457066	-9.797143	-2.6609352	-26.235407	0.0000	-0.14143	0.077621
Maximum	384.80725	12.3413111	4.40234578	1.32921919	11.34865	2.028797	103.5244
Mean	14.923527	0.93770849	0.97293327	0.7777418	1.260246	0.036056	2.160997

Table 3. Descriptive Statistics of the Beneish M-Score Ratio of the Fraudulent Group

Deskriptif	DSRI	GMI	AQI	SGI	DEPI	TATA	LVGI
Median	1.0268283	0.95777408	0.99913228	1.03920914	0.950902	-0.00175	0.985749
Std. Dev	0.1999347	1.27961319	0.50569703	2.94202187	0.457352	0.071815	1.974839
Minimum	-0.028703	-9.797143	-2.6609352	-26.235407	0	-0.14143	0.077621
Maximum	1.4533804	1.47373243	2.762601	1.32873419	3.943805	0.381949	12.47541
Mean	1.0045426	0.76988235	0.98618236	0.7159543	1.045949	-0.00237	1.362823

Table 4. Descriptive Statistics of the Beneish M-Score Ratio of the Fraudulent Group

Deskriptif	DSRI	GMI	AQI	SGI	DEPI	TATA	LVGI
Median	11.800627	0.97255279	0.99990684	1.06688881	0.950625	0.016327	0.989185
Std. Dev	97.060399	2.19187051	0.9689158	0.42750002	2.699566	0.471591	19.38567
Minimum	-1.457066	-0.147677	-1.6318705	-0.8058242	0.709808	-0.09745	0.124743
Maximum	384.80725	12.3413111	4.40234578	1.32921919	11.34865	2.028797	103.5244
Mean	57.674695	1.45317448	0.93223963	0.96751772	1.918443	0.154064	4.612533

Discriminant analysis of the ratios of the benefit scores in classifying fraudulent and non-fraudulent companies

1. Group mean similarity test

This test uses the Wilks' lambda value approach and the F test's significant value to determine the group mean similarity through its significance level. If the Wilks 'lambda value approaches zero (0) or the Significance value in the F test is less than 0.05, it indicates that it is more significant. The Wilks' lambda value approaches the number one (1), then it is not significant. Details of the Wilks' lambda value and the group average similarity test are presented in the following table:

Table 5. Tests of Equality of Group Means

Variables	Wilks' Lambda	F	df_1	df_2	Sig.
DSRI	0,893	13,413	1	112	0,000
GMI	0,970	3,517	1	112	0,063
AQI	0,963	4,285	1	112	0,041
SGI	1,000	0,000	1	112	0,988
SGAI	0,999	0,116	1	112	0,734
DEPI	0,867	17,119	1	112	0,000
TATA	0,933	8,053	1	112	0,005
LVGI	0,994	0,623	1	112	0,432

Based on table 4, it is found that the variables that differentiate between the manipulator and non-manipulator groups are DSRI, AQI, DEPI, and TATA because each of these variables has a significant value less than 0.05 and the wilks'lambda value is close to 0.

2. Significance test between the two variables

This test is used to find the best variable that differentiates it using the stepwise method. Table 6 shows the results of the variable significance test using the stepwise approach.

Tabel 6. Standardized Canonical Discriminant Function Coefficients

Variabel	Koefisien Canonical	**Sig.**
DSRI	0,674	0,000
GMI	0,355	0,000
AQI	0,396	0,000
DEPI	0,671	0,000
TATA	0,446	0,000

Based on Table 6, the efficiency of the DSRI, GMI, AQI, DEPI, and TATA variables is all significant, seen from the sig value which is less than 0.005.

Table 6 also shows that the discriminant function's value on the DSRI is 0.727 higher than the value of other variables. That means that DSRI is the most dominant variable in forming the discriminant function.

3. Discriminant Model Accuracy Test

The discriminant model accuracy test aims to determine how much the discrimination model obtained can explain the differences between companies classified as fraud and the group of companies that are not a fraud.

Table 7. Discriminant Model Accuracy Test

Function	**Eigenvalue**	**Wilks Lamda**	**Sig.**	**Canonical Correlation**
1	$0,571^{a}$	0,637	0,000	0,603

Based on table 7, the cannonical correlation value is 0.603. If it is squared (square canonical correlation) of 0.363, it means that the discriminant function can explain 36.3% of a manipulator and non-manipulator companies' variance. Moreover, when viewed from the significant value of the Wilk's lambda of 0,000, it means that between the two groups of a manipulator and non-manipulator companies, there is a significant difference from the discriminant function.

4. Establishment of Discriminant Functions

Table 8 shows each variable's coefficient from the discriminant analysis, which can then be used to form the discriminant function.

Table 8. Canonical Discriminant Function Coefficient

Variabel	Koefisien
DSRI	3,695
GMI	0,671
AQI	0,739
DEPI	0,568
TATA	6,472
(Constant)	-5,910

The discriminant functions formed based on table 9 are as follows:

D = -5,910 + 3,695 DSRI + 0,0671 GMI + 0,739 AQI + 0,568 DEPI + 6,472 TATA

5. Accuracy of Discriminant Function Classification

Tabel 9. Accuracy of Discriminant Function Classification

		Cuttoff_Beneish	Predicted Group Membership		Total
			0	1	
Original	Count	0	96	7	103
		1	5	6	11
	%	0	93,2	6,8	100,0
		1	45,5	54,5	100,0

a. 89,5% of original grouped cases correctly classified.

Table 9 shows that out of 103 non-manipulator companies, seven companies moved into the manipulator group. Meanwhile, of the 11 manipulator companies, five companies moved into the non-manipulator group. It can be concluded that the accuracy of the discriminant function in classifying manipulator and non-manipulator companies is 89.5%.

Discussion.

Based on the discriminant analysis of 8 Beneish M-Score ratios, there are only five significant ratios in grouping companies into manipulator and non-manipulator groups, namely DSRI, GMI, AQI, DEPI, and TATA. By looking at the significance test's statistical value, the DSRI value is 86.7% with a sig value. 0,000; GMI of 76.7% with a sig. 0,000; AQI of 70.4% with a sig. 0,000; DEPI is 66.7% with a sig. 0,000 and TATA of 63.7% with a sig. 0,000.

The discriminant analysis shows that DSRI is the most dominant ratio in the grouping of fraud companies and non-fraud companies (canonical discriminant value of 0.674). DSRI is a ratio that considers changes in the company's receivables with sales in a period with the previous period. The sample in this study is a banking sector company where the accounts receivable repair company or, in this case, is a credit bill that is closely related to sales or sales of the company, namely banking companies in the form of interest on credits/claims given to customers.

The discriminant function formed is D = -5,910 + 3,695 DSRI + 0,0671 GMI + 0,739 AQI + 0,568 DEPI + 6,472 TATA. This function proves that the company correctly classifies the companies into manipulator and non-manipulator companies by 89.5%. A total sample of 114 companies, according to Beneish M-Score, are divided into two groups, namely 103 non-manipulator companies and 11 manipulator companies. The 11 companies classified as manipulators indicated that fraud had occurred in their financial statements, so there is a need for further detection and prevention of fraud.

Conclusion

Based on the analysis and discussion results, if it is related to the signaling theory in the implications of this study, the signal theory will state that the role of encouragement played by managers, especially when submitting annual reports. Managers tend to provide good signals to stakeholders through financial reports. Managers often carry out even acts of fraud or manipulation of financial statements to meet related parties' expectations regarding company performance.

The results of the analysis related to the reliability of the beneish m-score in the detection of financial statement fraud can be concluded as follows:

1. The Beneish M-Score can detect fraud by 89.5%, meaning that this value is in the high category. It can be said that the Beneish M-Score model is accurate in detecting fraud in the financial statements of banking companies for the period 2016-2018.
2. The Beneish ratios that proved significant in the grouping of financial statements for the 2016-2018 period into the manipulator group of fraud companies and non-manipulator companies were DSRI, GMI, AQI, DPI, and TATA.
3. The ratio that differentiates or is the most dominant in grouping manipulator and non-manipulator companies is DSRI.

Based on the study results, it is recommended that investors and potential investors conduct financial analysis before investing because the presented financial reports often do not shows the company's real condition, so they must consider several aspects of the assessment. Furthermore, for further research, it is hoped that this research can become a reference material and expand the research object on banking companies that are declared to have committed fraud by Otoritas Jasa Keuangan.

References

M. D. Beneish, "Earnings management: A perspective," *Manag. Financ.*, vol. 27, no. 12, pp. 3–17, 2001.

R. Kassem and A. Higson, "The new fraud triangle model," *J. Emerg. Trend Econ. Manag. Sci.*, vol. 3, no. 3, pp. 191–195, 2012.

M. D. Beneish, "The Detection of Earnings Manipulation," *Financ. Anal. J.*, vol. 5, no. June, pp. 24–36, 1999.

Tarjo and N. Herawati, "Application of Beneish M-Score Models and Data Mining to Detect Financial Fraud," *Procedia - Soc. Behav. Sci.*, vol. 211, no. September, pp. 924–930, 2015.

M. Mahama, "Detecting corporate fraud and financial distress using the Altman and Beneish models," *Int. J. Econ. Commer. Manag.*, vol. 3, no. 1, pp. 1–18, 2015.

N. Omar, R. K. Koya, Z. M. Sanusi, and N. A. Shafie, "Financial Statement Fraud: a Case Examination sing Beneish Model and Ratio Analysis," *Int. J. Trade, Econ. Financ.*, vol. 5, no. 2, p. 184, 2014.

R. N. Kartikasari and G. Irianto, "Penerapan Model Beneish (1999) dan Model Altman (2000) dalam Pendeteksian Kecurangan Laporan Keuangan," *J. Akunt. Multiparadigma*, vol. 1, no. 2, pp. 323–340, 2010.

A. Mehta and G. Bhavani, "Application of forensic tools to detect fraud: the case of Toshiba," *J. Forensic Investig. Account.*, vol. 9, no. 1, pp. 692–710, 2017.

G. Bhavani and C. T. Amponsah, "M-Score and Z-Score for Detection of Accounting Fraud," *Account. Bus. Public Interes.*, vol. 3, no. 1, pp. 68–86, 2017.

K. Amoa-Gyarteng, "Analyzing a listed firm in Ghana for Early Warning Signs of Bankruptcy and Financial Statement Fraud: An Empirical Investigation of AngloGold Ashanti," *Eur. J. Bus. Manag.*, vol. 6, no. 5, pp. 10–17, 2014.

S. Santosa and J. Ginting, "Evaluasi Keakuratan Model Beneish M-Score Sebagai Alat Deteksi Kecurangan Laporan Keuangan (Kasus Perusahaan pada Otoritas Jasa Keuangan di Indonesia)," *Maj. Ilm. Bijak*, vol. 16, no. 2, pp. 75–84, 2019.

A. A. Arens, R. J. Elder, and M. Beasly, *Auditing and Assurance Services: an Integrated Approach*. 2012.

W. Allbrecht, C. Allbrecht, C. Allbrecht, and M. Zimbelman, *Fraud examination*. Cengage Learning, 2011.

M. D. Beneish, M. L. Charless, and D. C. Nichols, "No Title," *Financ. Anal. J.*, vol. 69, no. 2, pp. 57–82, 2013.