*European Research Studies Journal*
*Volume XXIII, Special Issue 2, 2020*

*pp. 622-637*

# Identification of a Fraudulent Organizational Culture in Enterprises Listed in Warsaw Stock Exchange

*Submitted 12/09/20, 1st revision 01/10/20, 2nd revision 15/10/20, accepted 22/11/20*

Joanna Wyrobek[1], Łukasz Popławski[2],
Marcin Surówka[3]

***Abstract:***

***Purpose:*** *This publication aims to verify the research hypothesis that it is possible to create a model predicting a fraudulent organizational culture in an enterprise.*
***Design/Methodology/Approach:*** *We estimated a logit model that would warn against fraud by employees, fraudulent financial reporting, and direct manipulation of the financial statements. The model would only caution about the possibility of the fraud risk and force a human analysis of such an entity.*
***Findings:*** *We searched for scams in the press and police investigations that ended with charges against directors, managers, and employees of enterprises where there was financial damage on a large scale. We were mainly looking for intentional and deliberate frauds.*
***Practtical Implications:*** *We created a model that warns which companies may have an unfair corporate culture for Polish companies listed on the WSE.*
***Originality/Value:*** *An element of the novelty in the publication is an attempt of a comprehensive approach to fraud, i.e., going beyond manipulating the financial result or penalties by the commission supervising the stock exchanges and financial markets.It is also one of the first fraud detection models which was designed for companies listed on the Warsaw Stock Exchange.*

***Keywords:*** *Fraudulent organizational culture, Warsaw Stock Exchange, fraud detection, big data.*

***JEL classification:*** *K40.*

***Paper Type:*** *Research study.*

***Acknowledgement:*** *The publication is financed by the Ministry of Science and Higher Education of Poland within ―Regional Initiative of Excellence‖ Programme for 2019-2022. Project no.: 021/RID/2018/19. Total financing: 11 897 131,40 PLN.*

[1] *Uniwersytet Ekonomiczny w Krakowie, wyrobekj@uek.krakow.pl;*
[2] *Uniwersytet Ekonomiczny w Krakowie, lukasz.poplawski@uek.krakow.pl;*
[3] *Uniwersytet Ekonomiczny w Krakowiea, marcin.surowka@uek.krakow.pl;*

## 1. Introduction

Business fraud results from the deliberate act of a man whose moral backbone has proved too weak to overcome opportunity, pressure, and the urge to rationalize wrongdoing. Business fraud usually does not kill anyone, but it destroys investor confidence in the capital market and institutions designed to keep them safe. Any statistics show that it is challenging to detect fraud without a whistleblower or denouncing a co-worker. A tiny percentage of crimes are detected by internal control, internal and external audits, or even the police and other control and supervision services.

This article aims to answer whether (since there are models for detecting fraudulent financial statements) is possible to create a model with a broader scope of action, identifying an unhealthy organizational culture in a company conducive to committing all kinds of fraud. If such circumstances occur, the moral decline should start from small cases to large issues. Large issues should affect the financial statements from some point because the financial statements are the result of practically all activities carried out by the enterprise (sooner or later, fraud leads to some entry in the accounting system units). In particular, the article will verify two research hypotheses. The first assumes that the models identify manipulation of financial results that can also indicate other purposeful activities in a criminal enterprise. The second hypothesis is that you can build a model that recognizes companies with an unethical (fraud-friendly) organizational culture.

The article will first briefly discuss the existing models of detecting fraudulent financial statements. They provide knowledge about variables and indicators that are good predictors of fraud in an enterprise. Then, we will explain the research method, and finally, the results of our research and the resulting conclusions.

## 2. Literature Review

Table 1 summarizes selected models for identifying the manipulation of financial statements. Most authors created models using various algorithms, machine learning, and artificial intelligence (some of them may also be called econometric methods). For the sake of brevity, Table 1 shows, however, for each of these publications, only one, the best model. We also only showed publications where accuracy and recall were above 50% or close to this limit.

As shown in Table 1, there are many publications dealing with the detection of financial statements fraud. Most of them use data from the US Financial Supervision Authority and Stock Exchanges, particularly Accounting and Auditing Enforcement Releases (AAERs). AAERs are Securities and Exchange Commission messages describing violations committed by companies, including intentional falsification of financial statements, over-statement of assets and income, and inadequate disclosure.

***Table 1.*** *Previous research results*

| Date | Accuracy | Recall | Algorithm | Source |
|---|---|---|---|---|
| Financial ratios | 47.6 | 63 | Logistic regression | Green, Calderon (1995) |
| Financial ratios | 71.25 | 72.52 | ANN | Green, Choi (1997) |
| Financial ratios | 83.54 | 81.08 | Logistic regression | Bell, Carcello (2000) |
| Financial ratios | 62.50 | 66 | ANN | Fanning, Cogger (1998) |
| Financial ratios, insider trading factors | 66.7 | 72.2 | Logistic regression | Summers, Sweeney (1998) |
| Financial ratios | 89.5* | 54.2* | unweighted probit | Beneish (1999) |
| Financial ratios | 69.72 | 81.03 | ANN | Feroz et al (2000) |
| Financial ratios | 87.75 | 86.29 | UTADIS | Spathis et al (2002) |
| Financial ratios | 76 | 35 | FNN | Lin et al (2003) |
| Financial ratios | 95.1 | 90.2 | Stacking | Kotsiantiset et al. (2006) |
| Financial ratios | 90.3 | 91.7 | Bayesian Belief Network | Kirkos et al. (2007) |
| Financial ratios | 95 | 63 | Genetic algorithms | Hoogs et al. (2007) |
| Financial ratios | ON | 98.39 | CART | Bai et al. (2008) |
| Financial + non-financial variables | 89.02 | on | UTADIS | Gaganis (2009) |
| Financial ratios | 67 | 81.3 | Three-phase cutting plane | Dikmen, Küçükkocaoğlu (2010) |
| Financial ratios | 90.4 | 80 | SVM | Cecchini et al. (2010a) |
| Onthology | 75.41 | 76.53 | NLP text recognition | Cecchini et al. (2010b) |
| Financial ratios | 67.3 | 68 | C4.5 (decision tree) | Humphreys et al (2011) |
| Financial ratios | 98.09 | 98.09 | Probabilistic neural network | Ravisankar et al. (2011) |
| Financial ratios | 64.41 | 65.59 | logistic regression | Dechow et al (2011) |
| Analyst calls + Financial variables | 89.03 | 24.69 | logistic regression | Larcker and Zakolyukina (2012) |
| Traits extracted from oral speech | on | on | logistic regression | Hobson et al. (2012) |
| Management report | 82.95 | 80.71 | logistic regression | Purda and Skillicom (2015) |
| Traits extracted from financial social media + financial variables + management reports | 80 | 83.04 | SVM | Dong et al. (2018) |

***Note:*** ** We chose the most popular version of the Beneish model with the proportion of costs for errors 40: 1 and unweighted probit.*
***Source:*** *Own study.*

These publications provide valuable insight into which financial indicators and data sources (financial statements, non-financial statements, management statements, online opinions, etc.) have proved most useful in creating models to detect companies under SEC investigations. Many of the models presented in Table 1 are too complex to be easily applied in daily work, but some of them are single-equation linear models, similar in concept to Altman's z-score model.

Probably the most popular and relatively simple models detect manipulation of financial results of companies is the Beneish m-score model (1999) and the Dechow *et al.* F-Score model (2011). These models are recognized for their high efficiency and the fact that they are single-equation linear models that are simple to be used by practitioners. These models were designed only to detect intentional manipulations of financial statements, for example, wrong recognition of costs and revenues, improper depreciation of assets, incorrect posting of transactions aimed to hide losses or to increase profits. We wanted to see whether they can detect other frauds such as winning bids for bribes, paying high wages for false patents and rationalizations, exploiting suppliers, cheating of investors and clients, loans to bankrupt companies, purchases of useless assets for personal gains for directors, etc. The Beneish m-score model has the following form and it is presented in Table 2:

mScore = -4.84 + 0.92 * DSRI + 0.528 * GMI + 0.404 * AQI + 0.892 * SGI + 0.115 * DEPI - 0.172 * SGAI + 4.679 * TATA -0.327 * LVGI

***Table 2.*** *Definitions of indicators in the m-Score Beneish model*

| variable | definition |
|---|---|
| DSRI (days' sales in receivables index) | [Net receivables (t) / Sales (t)] / [Net receivables (t-1) / Sales (t-1)] |
| GMI (gross margin index) | [Sales (t-1) - Sales Costs (t-1)] / Sales (t-1) / [Sales (t) - Sales Costs (t)] / Sales (t) |
| AQI (asset quality index) | [1 - {Current Assets (t) + Property Plant and Equipment (t) + Securities (t)} / Total Assets (t)] / [1 - {Current Assets (t-1) + Property Plant and Equipment (t-1) + Securities (t-1) } / Total Assets (t-1)] |
| SGI (sales growth index) | Sales (t) / Sales (t-1) |
| DEPI (depreciation index) | Depreciation(t-1) / [PP&E (t-1) + Depreciation(t-1)] / Depreciation(t) / [PP&E(t) + Depreciation(t)] |
| SGAI (sales, general, and administrative expenses index) | [SellingGeneralAdministratCosts (t) / Sales (t)] / [SellingGeneralAdministratCosts (t-1) / Sales (t-1)] |
| LVGI (leverage index) | [Current liabilities (t) + Long-term debt (t)] / Total assets (t) / [Current liabilities (t-1) + Long-term debt (t-1)] / Total assets (t-1) |
| TATA (total accruals to total assets) | [Income from continuing operations (t) - OperatingCashFlows (t)] / Total Assets (t) |

***Source:*** *Own study based on Beneish (1999).*

The simplest F-score model by Dechova *et al.* (2011) has the following form and it is presented in Table 3:

FScore = -7.893 + 0.79 * RSSTAccruals + 2.518 * ChangeReceivables + 1.191 * ChangeInventories + 1.979 * PerctChangeSoftAssets + 0.171 * ChangeCashSales - 0.932 * ChangeROA + 1.029 * ActualIssuance

***Table 3.*** *Definitions of indicators in Dechova's F-Score model*

| variable | definition | Additional explanations |
|---|---|---|
| RSSTAccruals | (KON change + NCO change + FIN change) / Average total assets | KON = Current assets - Cash and short-term investments - (Current liabilities - Debt in current liabilities)), NCO = [Total assets - Current assets - Long-term investments and advances] - [total liabilities - current liabilities - long-term debt], FIN = [Short-term investments + Long-term investments] - [Long-term debt + Debt in current liabilities + Preference shares]. |
| ChangeReceivables | Change in receivables = [receivables (t) - receivables (t-1)] / [[total assets (t) + total assets (t-1)] / 2] | |
| ChangeInventories | Change in inventories = [inventories (t) - inventories (t-1)] / [[total assets (t) + total assets (t-1)] / 2] | |
| PerctChangeSoftAssets | % Soft assets = [soft assets (t) - soft assets (t-1)] / soft assets (t-1) | soft assets = (Total Assets - Tangible Fixed Assets - Cash and Cash Equivalents) / Total Assets |
| ChangeCashSales | Change of cash sales = (cash sales(t) - cash sales (t-1)) / cash sales (t-1) | cash sales = sales revenues - (receivables (t) - receivables (t-1)) |
| ChangeROA | Change in return on assets = [Net profit or loss (t) / average total assets (t)] - [net profit or loss (t-1) / average total assets (t-1 )] | |

| variable | definition | Additional explanations |
|---|---|---|
| ActualIssuance | Actual issuance = a discrete variable with a value of 1 if the company has issued ordinary or preference shares in a given year, and a value of 0 in other cases | |

***Source:*** *Own study based on Dechowa et al. (2011).*

The Beneish m-score model suggests manipulation of the financial statement if the m-score exceeds the value of -2.22. The Dechowa's F-Score model requires the calculation of the index, and then substitution into the formula for calculating the probability: P = exp (index) / (1 + exp (index)). The resulting probability is divided by the overall probability of fraud in a given population of enterprises - in Dechowa's paper; it was 0.0037: F-Score = P / 0.0037. The obtained result shows how many times a given enterprise has a greater probability of falsifying financial statements than a randomly selected enterprise from the entire surveyed population.

## 3. Methodology

In the paper, we wanted to check to what extent financial statement manipulation models (Beneish and Dechowa's) can detect unethical organizational culture (that is, identify companies in which business fraud occurs, which is a broader category than financial statement manipulation) and create our model dedicated to fraud detection in Polish companies.

Firstly, we tested whether popular financial statement manipulation models (Beneish and Dechowa's models) are capable of detecting a broader category of financial fraud than only statement manipulation (bribery, extortions, exploitation of suppliers, cheating of clients, false investments, misappropriation of real estate, sales of shares at a reduced price, etc.).

To test the Beneish and Dechowa's models, we applied them to the WSE (Warsaw Stock Exchange) companies, which committed fraudulent behavior. We took financial statements from the Notoria Service database. Fraud information came from 3 sources. Type A irregularities concerned the Police investigations with charges against companies or their employees. We took such information from the press, the Internet, and the Reuters database. Type B irregularities concerned companies penalized by the Polish Financial Supervision Commission (pl. KNF – Komisja Nadzoru Finansowego) mostly for improper preparation of financial statements. Type C irregularities represented situations where the auditor refused to issue an opinion or issued a negative opinion (either for the annual or semi-annual report, (Golec, 2019)). Table A.1 in the appendix provides a brief description of the fraud cases used in our paper.

For holding companies, we used consolidated yearly financial reports, and for the rest of the companies, we used individual financial statements. For each company, we selected at least one sister-company without fraudulent behavior. We matched sister companies based on the size and activity type. We thoroughly checked all sister companies for not experiencing any frauds. In total, we collected data for 93 fraudulent companies and 144 non-fraudulent companies. We collected in total 441 annual financial statements. For type A frauds, due to the unfinished court cases (even if the court case was resolved, convicts could appeal against the ruling), the company names were replaced with letters. For all companies presented in Table A.1, we calculated the Beneish and Dechowa models (for the years in which the irregularities occurred), and we checked their effectiveness in their detection.

In the second part of the research, we created our model for identifying companies with unethical organizational culture. We used for this purpose, a logistic regression model. We divided the research sample into two parts with proportions 80:20. We used the first part to train the models (with 10-fold cross-validation), and the remaining 20% of the companies served as a validation sample. Formula 1 presents the function called sigmoid, which we used to estimate a model.

$$p(y=+1|x_i,\mathbf{w}) = g\left(\mathbf{w}^T h(x_i)\right) = \frac{1}{1+e^{-Score(x_i)}} = \frac{1}{1+e^{w^T h(x_i)}} = \frac{1}{1+e^{-(w_0 h_0(x_i)+\cdots w_n h_n(x_i))}} \tag{1}$$

where: **w** – parameter vector, h – variable function x(i) (often h(x) = x).
The function estimation uses the transformed version of formula 2:

$$ln\left(\frac{p}{1-p}\right) = w_0 h_0(x_i) + w_1 h_1(x_i) + \cdots + w_n h_n(x_i) \tag{2}$$

where: p – probability, that observation x(i) belongs to class y = 1.
The optimization condition for determining the vector **w** of function parameters is given by equation 3. Logistic regression attempts to find such **w** parameters for which the logarithm of probability is the highest for all available instances.

$$\underset{max}{ln(\ell(w))} = max\left(\sum ln(p(y=+1|\boldsymbol{x},\boldsymbol{w})) + \sum ln(p(y=-1|\boldsymbol{x},\boldsymbol{w}))\right) \tag{3}$$

where: $\ell(w)$ – the first derivative of the likelihood function, **w** – parameter vector, **x** – observation vector.

Since there are two possible values of y (because the classification assigns observations to one of two classes, y = 0 or y = 1), then if the sum of probabilities in one class is maximized, and consequently, the sum of probabilities in the second class is automatically minimized. For this reason, it is necessary to specify a threshold for maximization. The default value is 0,5 for both sums, and we used this threshold for our model.

Parameter estimation was based on maximizing the probability function. If it is assumed that x(i) represents random samples taken from the f(**x**|**w**) distribution, the latter represents the probability distribution function (based on the assumption that the x distribution is a continuous distribution). With this assumption, for each sample of empirical observations (each instance can be a vector), the likelihood function is as follows:

$$\ell(w) = f(x_1, \dots, x_n|w) = f(x_1|w) \dots f(x_n|w) \tag{4}$$

Which, in turn, can be written as:

$$\ell(w: x_1, x_2, \dots, x_n) = p(x_1, x_2, \dots, x_n|w) = p(x_1|w)p(x_2|w) \dots p(x_n|w) \tag{5}$$

The last step in the estimation is to apply the natural logarithm equation to both sides to get rid of the product of probabilities.

$$\ell\ell(w) = ln \prod_{i=1}^{n} p(y_i|x_i, w) \tag{6}$$

The logit regression parameters presented in the finding section represent the inverse of the logistic regression model (F):

$$g(F(\boldsymbol{x})) = ln\left(\frac{F(\boldsymbol{x})}{1 - F(\boldsymbol{x})}\right) = w_0 + \boldsymbol{wx} \tag{7}$$

Independent variables used to create the model could not have a correlation coefficient higher than 0.4.

## 4. Research Results and Discussion

Table 4 shows the performance of the Beneish model broken down into abnormalities of type: A, B, and C. Table 5 shows the efficiency of the Dechowa model.

***Table 4.*** *The efficiency of the m-Score Beneish model on the sample of Polish companies listed on the WSE*

| Category | Empirical: Manipulator | Empirical: Non-manipulator | Error [%] | parameters | value |
|---|---|---|---|---|---|
| TOTAL group A, B, C | Manipulator | Non-manipulator | | accuracy | 67.7% |
| Model : manipulator | 36 | 48 | 1st type : 21.6 | precision | 42.9% |
| Model : non-manipulator | 52 | 174 | 2nd type : 59.1 | recall | 40.9% |
| GROUP A | Manipulator | Non-manipulator | | accuracy | 75.0% |
| Model : manipulator | 10 | 15 | 1st type : 12.9 | precision | 40.0% |
| Model : non-manipulator | 22 | 101 | 2nd type : 68.75 | recall | 31.3% |
| GROUP B | Manipulator | Non-manipulator | | accuracy | 53.4% |
| Model : manipulator | 7 | 22 | 1st type: 40.1 | precision | 24.1% |
| Model : non-manipulator | 12 | 32 | 2nd type: 63.16 | recall | 36.8% |
| GROUP C. | Manipulator | Non-manipulator | | accuracy | 67.4% |
| Model : manipulator | 19 | 11 | 1st type: 21.2 | precision | 63.3% |
| Model : non-manipulator | 18 | 41 | 2nd type: 51.35 | recall | 51.4% |

***Source:*** *Own study.*

***Table 5.*** *The efficiency of the F-Score Dechowa model on the sample of Polish companies listed on the WSE*

| Category | Empirically: Manipulator | Empirical: Non-manipulator | Error [%] | parameter | value |
|---|---|---|---|---|---|
| TOTAL group A, B, C | Manipulator | Non-manipulator | | accuracy | 64.7% |
| Model : manipulator | 14 | 17 | 1st: 5.4 | precision | 45.2% |
| Model : non-manipulator | 153 | 298 | 2nd type : 91.6 | recall | 8.4% |
| GROUP A | Manipulator | Non-manipulator | | accuracy | 68.6% |
| Model : manipulator | 2 | 5 | 1st type : 3.67 | precision | 28.6% |
| Model : non-manipulator | 56 | 131 | 2nd type : 94.9 | recall | 3.4% |
| GROUP B | Manipulator | Non-manipulator | | accuracy | 68.0% |
| Model : manipulator | 6 | 9 | 1st type: 8.7 | precision | 40.0% |
| Model : non-manipulator | 38 | 94 | 2nd type: 13.6 | recall | 13.6% |
| GROUP C | Manipulator | Non-manipulator | | accuracy | 56.0% |
| Model : | 6 | 3 | 1st type: 3.9 | precision | 66.7% |

| Category | Empirically: Manipulator | Empirical: Non-manipulator | Error [%] | parameter | value |
|---|---|---|---|---|---|
| manipulator | | | | | |
| Model : non-manipulator | 59 | 73 | 2nd type: 90.7 | recall | 9.2% |

***Source:*** *Own study.*

As can be seen from Tables 4 and 5, the overall efficiency of both models for Polish data was not impressive. Overall accuracy for the Beneish model was 67.7%, and for the Dechowa model, 64.7%. Models (especially of Dechowa) identified fair enterprises more correctly than unfair firms. The effectiveness of both models was similar. Even though the models' efficiency was not impressive, still they classified at least half of the companies in the correct class: A, B, or C (group A represented companies with police charges for financial crime, group B included companies with penalties enforced by the Polish Financial Supervision Commission and group C included cases of evident statement manipulation detected by auditors).

In the last stage of the research, we created our model for identifying the companies in which the irregularities occurred. Table 6 presents the model parameters, and Table 7 shows the model efficiency ratios calculated for the entire sample and the three types of fraud. A total of 441 observations were used in the study.

***Table 6.*** *Parameters of the proprietary logistic regression model for detecting A, B, C type frauds (unweighted logit)*

| Manipulator | Coef. | Std.Err. | z | P> \| z \| | [95% Conf. Interval] | |
|---|---|---|---|---|---|---|
| Intangible assets/Total Assets | 6.554414 | 1.672336 | 3.92 | 0 | 3.276695 | 9.832133 |
| Long-term provision for employee/Total Assets | 27.21811 | 7.593389 | 3.58 | 0 | 12.33534 | 42.10088 |
| Purchase of property plant and equipment/Sales revenues | 5.4271 | 1.522182 | 3.57 | 0 | 2.443983 | 8,410826 |
| Gross profit / loss on sales | -1.03948 | 0.463974 | -2.24 | 0.025 | -1.94885 | -0.13011 |
| Altman z-score (Altman, 1968, Table A.2*) | -0.08873 | 0.021845 | -4.06 | 0 | -0.13155 | -0.00664 |
| Mączyńska score (Mączyńska 2004, Table A.3**) | -0.01418 | 0.003848 | -3.68 | 0 | -0.21722 | -0.00664 |
| constant | -0.37662 | 0.176912 | -2.13 | 0.033 | -0.72336 | -0.02988 |

Where: * Altman z-score model is described in Appendix, Table A.2, ** Mączyńska model is described in Appendix, Table A.3.

***Table 7.*** *The efficiency of own logistic regression model used to detect A, B, C fraud (validation sample, unweighted logit)*

| Category | Empirically: Manipulator | Empirical: Non-manipulator | Error [%] | parameter | value |
|---|---|---|---|---|---|
| TOTAL group A, B, C | Manipulator | Non-manipulator | | accuracy | 75.28% |
| Model : manipulator | 9 | 2 | 1 type: 3.3 | precision | 81.82% |
| Model : non-manipulator | 20 | 58 | 2 type: 68.97 | recall | 31.03% |
| GROUP A | manipulator | Non-manipulator | | accuracy | 72.97% |
| Model : manipulator | 3 | 3 | 1 type: 11.1 | precision | 50.00% |
| Model : non-manipulator | 7 | 24 | 2 type: 70 | recall | 30.00% |
| GROUP B | manipulator | Non-manipulator | | accuracy | 73.08% |
| Model : manipulator | 1 | 0 | 1 type: 0 | precision | 100.00% |
| Model : non-manipulator | 7 | 18 | 2 type: 87.5 | recall | 12.50% |
| GROUP C | manipulator | Non-manipulator | | accuracy | 92.31% |
| Model : manipulator | 10 | 1 | 1 type: 6.7 | precision | 90.91% |
| Model : non-manipulator | 1 | 14 | 2 type: 9.1 | recall | 90.91% |

To use our model one should firstly calculate the linear combination of products between parameters and variables:

$w^T h(x_i)$ = 6.5544 * Intangible Assets/Total Assets + 27.2181 * Long-term provisions for employees/Total Assets + 5.4271 * Purchase of property, plant and equipment/Total sales revenues – 1.0394 * Gross profit/loss on sales / total sales revenues – 0.08873 * Altman z-score – 0.01418 * Mączyńska model score – 0.37662 (8)

Then, one should substitute the result of formula 8 into the sigmoid function:

$$p(y=+1|x_i,\mathbf{w}) = \frac{1}{1+e^{w^T h(x_i)}} \tag{9}$$
$$= \frac{1}{1+e^{-\left(6.5544*\frac{IntA}{TA}+27.2181*\frac{LTPE}{TA}+5.4271*10^6*PPE/TS-1.0394*GPLS/TS-0.08873*Altman-0.01418*Maczynska-0.37662\right)}}$$

Obtained probability informs about the chances that in a particular year, a company could have committed fraud.

As shown in Table 7, our model of fraud detection has a higher efficiency than the analyzed models created by Beneish and Dechowa. The model structure suggests that the risk of fraud in the enterprise increases with the increase in intangible assets in total assets, and with the share of provisions for employees in total assets. Another important risk factor is the relation between purchases of tangible fixed assets and sales revenues.

The risk of fraud decreases with an increase in sales profitability, an increase in the z-score index, and Mączyńska's model index. This is quite an unexpected result, as both the Mączyńska and Altman models suggest no risk of bankruptcy and a good financial situation if the value of these models is high for a given enterprise. Thus, at least in the analyzed sample of enterprises, a tendency to fraud and manipulation was observed for entities for which there was no direct threat of bankruptcy (in a given year).

If we look at Table 7, the model showed the highest efficiency for group C, where accuracy, precision, and recall were over 90%. Thus, the years in which manipulating the financial statements took place, where the auditor refused to issue an opinion or issued an adverse opinion, are very well detected by the model. For the remaining groups of frauds, the effectiveness of the model was lower, but the model coped quite well with group A frauds, slightly worse with group B fraud. However, the model rarely makes a type I error, i.e., false alarms were rare.

## 5. Conclusions, Proposals, Recommendations

When one analyzes results of research on the identification of companies committing various frauds, it seems that in the analyzed sample, most fraudulent firms were characterized by good (in the meaning: in no direct threat of bankruptcy), but not the ideal financial situation (profitability decreased chances of fraud). Fraudulent companies were also characterized by massive expenditure on tangible fixed assets, high long-term provisions for employees, and a high share of intangibles in total assets. This relationship suggests a link between fraud and high investment expenses, where management is likely to focus on investments and pay less attention to other operating issues. High long-term provisions for employees suggest setting aside reserves for awards, severance payments, and profit payments for employees, which does not prevent fraud but encourages it.

We observed that in companies prone to fraud, the gross profitability on sales was not too high; the higher was the profitability, the lower was the likelihood of fraud - perhaps in very profitable and wealthy companies, employees were well remunerated, and companies could afford effective control systems.

Summing up, it seems that the Beneish and Dechowa models, especially when they are applied together, can prove useful in identifying evident manipulations of the financial statements (group C), situations when the auditor refuses to issue an opinion (group B), and when the company is under the police investigation of severe crimes (group A). However, we managed to create a more efficient model, probably better adjusted for the Polish companies. Our model performed better with all types of fraud that we considered, especially with the groups B and C.

As for the practical applications of our research, we believe that one can use our model as a warning system for market investors to signal the possibility of fraud and manipulation in a particular company in a given year. Fraud does not necessarily lead a company immediately into bankruptcy, but such a company can experience serious financial problems in the future. Signals from all the models that we considered in this publication can be regarded as signals of poor internal control in entities and warnings that further, more severe fraud may occur, as the lack of a penalty usually leads to even greater offenses.

## References:

Altman, E. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance, 23, 589-609.

Bai, J., Yen, B. 2008. False Financial Statements: Characteristics of China's Listed Companies and CART Detecting Approach. Int. J. Inf. Technol. Decis. Poppy., 7, 339-359.

Bell, T.B., Carcello, J.V. 2000. A decision aid for assessing the likelihood of fraudulent financial reporting. Auditing, 19, 168-184.

Beneish, M. 1999. The Detection of Earnings Manipulation. Financ. Anal. J., 5, 24-36.

Cecchini, M., Aytug, H., Koehler, G.J., Pathak, P. 2010. Detecting management fraud in public companies. Manage. Sci., 56, 1146-1160.

Cecchini, M., Aytug, H., Koehler, G.J., Pathak, P. 2010. Making words work: Using financial text as a predictor of financial events. Decis. Support Syst., 50, 164-175.

Dechow, P.M., Ge, W., Larson, C.R., Sloan, R.G. 2011. Predicting Material Accounting Misstatements. Contemp. Account. Res., 28, 17-82.

Dikmen, B., Ḱkkocaŏlu, G. 2010. The detection of earnings manipulation: The three-phase cutting plane algorithm using mathematical programming. J. Forecast., 29, 442-466.

Dong, W., Liao, S., Zhang, Z. 2018. Leveraging Financial Social Media Data for Corporate Fraud Detection. J. Manag. Inf. Syst., 35, 461-487.

Fanning, K.M. Cogger, K.O. 1998. Neural network detection of fraud management using published financial data. Int. J. Intell. Syst. Accounting, Financ. Manag., 7, 21-41.

Feroz, E.H., Kwon, T.M., Pastena, V.M., Park, K. 2000. The efficacy of red flags in predicting the SEC's targets: an artificial neural networks approach. Int. J. Intell. Syst. Accounting, Financ. Manag., 9, 145-157.

Gaganis, C. 2009. Classification techniques for the identification of falsified financial statements: a comparative analysis. Intell. Syst. Accounting, Financ. Manag., 16, 207-229.

Golec, A. 2019. Ocena skuteczności modelu Beneisha w wykrywaniu manipulacji w sprawozdaniach finansowych. Folia Oeconomica, Acta Universitatis Lodziensis 2(341), 161-182.

Green, B.P. 1995. Analytical procedures, and the auditor's capacity to detect management fraud. Accounting Enq. A Res. J., August, 1-48.

Green, B.P., Choi, J.H. 1997. Assessing the risk of management fraud through neural network technology. Auditing, 16(1997) 25-28.

Hobson, J.L., Mayew, W.J., Venkatachalam, M. 2012. Analyzing Speech to Detect Financial Misreporting. J. Account. Res., 50, 349-392.

Hoogs, D., Kiehl, B., Lacomb, T., Senturk, C. 2007. A Genetic Algorithm Approach to Detecting Temporal Patterns Indicative of Financial Statement Fraud. Intell. Syst. Accounting, Financ. Manag., 15, 41-56.

Humpherys, S.L., Moffitt, K.C., Burns, M.B., Burgoon, J.K., Felix, W.F. 2011. Identification of fraudulent financial statements using linguistic credibility analysis. Decis. Support Syst. 50, 585-594.

Kirkos, E., Spathis, C., Manolopoulos, Y. 2007. Data Mining techniques for the detection of fraudulent financial statements. Expert Syst. Appl., 32, 995-1003.

Kotsiantis, S., Koumanakos, E., Tzelepis, D., Tampakas, V. 2006. Forecasting fraudulent financial statements using data mining. Int. J. Comput. Intell., 3, 104-110.

Larcker, D.F., Zakolyukina, A.A. 2012. Detecting Deceptive Discussions in Conference Calls. J. Account. Res., 50, 495-540.

Leuz, C., Nanda, D., Wysocki, P.D. 2003. Earnings management, and investor protection: An international comparison. J. Financ. Econ., 69, 505-527.

Lin, J.W., Hwang, M.I., Becker, J.D. 2003. A fuzzy neural network for assessing the risk of fraudulent financial reporting. Manag. Audit. J., 18, 657-665.

Mączyńska, E. 2004, Systemy wczesnego ostrzegania. Nowe Życie Gospodarcze, 12(373), 4-9.

Purda, L., Skillicorn, D. 2015. Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection. Contemp. Account. Res., 32, 1193-1223.

Ravisankar, P., Ravi, V., Raghava Rao, G., Bose, I. 2011. Detection of financial statement fraud, and feature selection using data mining techniques. Decis. Support Syst., 50, 491-500.

Spathis, C., Doumpos, M., Zopounidis, C. 2002. Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques. Eur. Account. Rev., 11, 509-535.

Summers, J.T., Sweeney, S.L. 1998. Fraudulently misstated financial statements and insider trading: An empirical analysis. Account. Rev., 73, 131-146.

**Appendix:**

***Table A.1.*** *Brief description of fraud cases used in the model training*

| Fraud type | Number of cases | Number of sister companies | Description |
|---|---|---|---|
| Type A | 28 | 34 | * fraud and misappropriation of real estate and company shares,<br>*sale of shares in a subsidiary at a reduced price,<br>* a series of deliberate omissions which led to the purchase of another company at an inflated price,<br>* accepting bribes for signing contracts, corruption,<br>* extortion of property by employees and management for false patents and rationalizations,<br>* speculation of Supervisory Board members on the share prices,<br>* falsifying financial statements and taking money out of the company,<br>* concealment of information about the company's situation when the company issued bonds,<br>* cheating customers,<br>* cheating contractors as to the actual situation of the company,<br>* economically unjustified purchase of an unnecessary investment that caused considerable financial losses for the company,<br>* an economically unjustified contract that exposed the company to large financial losses,<br>* granting loans to an entity in a terrible financial situation that did not give a chance to repay these loans,<br>* signing a fictitious contract for consulting services,<br>* improper supervision over sponsored events exposed the company to financial losses,<br>* leading to the signing of an unfavorable contract with the company,<br>* entering into contracts in return for bribes,<br>* dramatically underselling real estate sales and money laundering,<br>* tax offences,<br>* misrepresentation as to the actual situation of the enterprise at the time of issuing the bonds |
| Type B | 24 | 28 | improper preparation of the annual (or semi-annual) consolidated or non-consolidated financial statements |
| Type C | 41 | 57 | Refusal to issue an opinion by a statutory auditor or a negative opinion |
| Total | 93 | 119 | x |

***Table A.2.*** *Altman's z-score model parameters*

| indicator | Coef |
|---|---|
| working capital/total assets | 1.2 |
| retained earnings/total assets | 1.4 |
| earnings before interest and taxes/total assets | 3.3 |
| market value of equity/total liabilities | 0.6 |
| sales/total assets | 0.999 |
| X >= 2.99 - "safe zone" | |
| 2.00 >= X >= 1.81 - "grey zone" | |
| x <= 1.81 - "distress zone" | |

***Table A.3.*** *Mączyńska model parameters*

| indicator | Coef |
|---|---|
| operating profit/total assets | 9.498 |
| equity/total assets | 3.566 |
| (net result+depreciation)/ total liabilities | 2.903 |
| current assets/short-term liabilities | 0.452 |
| constant | -1.498 |
| X < 0 "bankrupt | |
| X >= 0 "non-bankrupt" | |