

Papík, Mário; Papíková, Lenka

Article

Detection models for unintentional financial restatements

Journal of Business Economics and Management (JBEM)

Provided in Cooperation with:

Vilnius Gediminas Technical University

Suggested Citation: Papík, Mário; Papíková, Lenka (2020) : Detection models for unintentional financial restatements, Journal of Business Economics and Management (JBEM), ISSN 2029-4433, Vilnius Gediminas Technical University, Vilnius, Vol. 21, Iss. 1, pp. 64-86, <https://doi.org/10.3846/jbem.2019.10179>

This Version is available at:

<https://hdl.handle.net/10419/317379>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

DETECTION MODELS FOR UNINTENTIONAL FINANCIAL RESTATEMENTS

Mário PAPÍK^{*}, Lenka PAPÍKOVÁ

*Department of Economics and Finance, Faculty of Management,
Comenius University, Bratislava, Slovakia*

Received 05 April 2018; accepted 10 April 2019

Abstract. The aim of manuscript is to analyze and identify determinants of honest accounting errors leading to financial restatements based on data from SEC database and from annual reports. Reason for this study is that accounting errors are expensive for companies that need to change already published financial statements and have impact on company reputation and stock price. Most of authors focus on prediction of accounting frauds and financial restatements remain in the background of research. This study initially tests existing accounting fraud detection model of Beneish on a sample of 40 financial restatement companies over 10 years and develops two new pioneer prediction models, one based on linear discriminant analysis (LDA) and another based on logistic regression. In testing dataset, LDA model has achieved accuracy 70.96%, specificity 25.00% and sensitivity 79.83% and logistic regression model has achieved accuracy 62.22%, specificity 41.66% and sensitivity 66.67%, performance of both models is better than existing Beneish model or other studies in this field. Developed models can be widely used by both internal and external users of financial statements, who would like to determine if financial statements of analyzed company include accounting errors or not, thanks to easily interpretable results in equation form.

Keywords: unintentional financial restatement, financial restatements, accounting fraud, accounting error, linear discriminant analysis, logistic regression, prediction modelling, fraudulent financial statements, accounting manipulation, auditing.

JEL Classification: C81, M41.

Introduction

Key aim of financial statements that are prepared meaningfully is that financial statements are comparable to financial statements of other companies. This is even more important for a solid and correct comparison of companies in the same industry. Meaningful comparison of companies across industries requires that financial statements are prepared in line with local Generally Accepted Accounting Principles.

*Corresponding author. E-mail: mario.papik@gmail.com

Consequences of poor corporate governance and financial reporting processes include consequences like bankruptcies, frauds, earnings restatements, quick and sudden decrease of corporate stock value, loss of credibility and business partners, loss of customers and many other negative impacts on the activities of companies (Savčuk, 2007).

There is a difference between accounting errors (leading up to financial restatements) and accounting frauds. The main difference is in the root cause of financial restatement or accounting fraud. It must be determined if an honest accounting, reporting or clerical mistake has occurred or if intentional misrepresentation or fraudulent reporting has been done. Situations occur when an error in financial reporting is discovered. In such a case, this error should be corrected, and one way of correction is to prepare financial restatement. For example, under US GAAP, a company must prepare financial restatement when discovered error is significant to prior period financial statements. Financial restatement is prepared in a Form 10-K/A filing – along with an audit opinion to the disclosure of restatement.

Documents from developed countries such as the United States of America or the European Union, indicate that joint actions concerning financial reporting and especially the creation of Code of Professional Ethics for international accounting companies and auditors, can be used in prevention of accounting kind of criminal activity (Bhasin, 2013; Sadaf, Oláh, Popp, & Mate, 2018). Similar findings have also been discovered in some studies considering Japan and occurrence of accounting fraud in Japan. These studies pointed out to positive correlation between high ethical standards of this country and high quality of its legislative framework and lower occurrence of accounting fraud (Nakashima, 2017; Song, Oshiro, & Shuto, 2016). Similarly, studies from less developed countries such as Nigeria, where the relationship between forensic accounting practices and fraud prevention has been also proved to have positive correlation (Oyebisi, Wisdom, Olusogo, & Ifeoluwa, 2018), or studies from Jordan, which determined that family businesses are more prone to earnings management than businesses without family members present in the board of directors (Bateineh, Abuadous, & Alabood, 2018).

Audit Analytics research (2015) analyzes financial restatements of companies as a financial reporting phenomenon over the period from 2004 to 2014. Research sample with 14,000 financial restatements, which included 9,000 companies with financial restatement, analyzes various parameters. This study identifies 10 major accounting items during analyzed period, in which companies generally restate their financial statements: (1) debt, quasi-debt, warranty and property; (2) change in cash; (3) tax expense, benefits and deferrals; (4) foreign, affiliate and subsidiary affairs; (5) remuneration, sales and administration expenses; (6) revenue recognition; (7) liabilities, reserves and accruals; (8) receivables, investments and cash; (9) inventories, costs of goods sold and (10) tangible and intangible assets and their evaluation. In 2014, the highest amount (23% of total) of financial restatements concerns corrections of the debt-related accounts, followed by cash accounts (20%) and tax expense accounts (14%). Financial restatements related to errors in revenue recognition account for 11% of the total number of cases analyzed.

In 2002, the highest financial restatement impact on a company net profit is \$4.5 billion, which is mainly due to an overestimation of company earnings. On the other hand, in 2014, the most significant financial restatement is worth only \$154 million. Despite decreasing trend in the amount and value of financial restatements, these restatements, unlike bank-

ruptcy and company failure that reflect previous company financial credit crisis, financial restatement can be considered a breach of trust towards stakeholders and especially investors (Chen & Zhou, 2010).

In addition, when an audited financial statements are restated, the validity of the auditor's opinion and the audit process will sometimes be questioned because the published data are, on occasion, not free of critical errors (Stanley & DeZoort, 2007). Due to negative impact of financial restatements on credibility of any company, this research has focused on creating models that could contribute to prediction of honest accounting errors leading to financial restatements (Gleason, Jenkins, & Johnson, 2008).

The aim of this manuscript is to analyze determinants which lead companies to honest accounting errors leading to financial restatements. To successfully fulfil the aim of this manuscript, existing accounting fraud prediction model – the Beneish model (Beneish et al., 1999) – has been tested and it has been determined whether this model could be also used to predict unintentional financial restatements, and with what attributes such as accuracy, sensitivity and specificity would this model perform.

Furthermore, new mathematical prediction models for unintentional financial restatements has been developed. So far, only one study (I. Dutta, S. Dutta, & Raahemi, 2017) is trying to develop own prediction model by using modern data mining methods such as ANN, Support vector machine and Bayesian Belief. Although these techniques are highly accurate, they do not allow users to create simple intuitive models as some classical techniques, which is known from Beneish's fraud model (1999) or well-known Altman's bankruptcy model (1968).

Due to this reason, two mathematical and statistical techniques – of logistic regression and linear discriminant analysis (LDA) – are used in this research. Application of logistic regression and LDA enables users to effectively detect potential financial restatements. For easy application in practical real-life examples, new developed models are presented in a form of equation which is easy to use and results are easy to interpret. Development of new prediction models is based on two input variable datasets – training and testing – both of which are described in section 2.1. in more detail. Performance of models developed in this study is subsequently determined by attributes: accuracy, sensitivity and specificity and subsequently compared to the Beneish M-score and Dutta's model. Since this manuscript is one of the first financial restatement prediction manuscripts, there are not any other prediction studies that could be compared to results of this study.

The paper is structured as follows: Section 1 contains literature review of relevant prior studies. Section 2 describes used research methodology. Section 3 provides list and description of developed models and results of performed analysis. Section 4 contains discussion and last section represents the concluding remarks.

1. Literature review

Accounting fraud in company financial statements and fraudulent reporting is a deliberate act by companies to deceive users of publicly available financial statements – especially investors and creditors – by preparing and publishing manipulated financial statements (Rezaee,

2005). The Enron accounting scandal has revealed three fundamental principles under Generally Accepted Accounting Principles (GAAP) that heralded its fall: (1) The off-balance sheet arrangements, (2) The role of mark-to-market, and (3) The manipulation of derivatives (Lemus, 2014). Large list of available foreign literature on the topic of fraudulent activities exists (R. Duska & B. Duska, 2011; Girgenti & Hedley, 2011; La Torre, 2009; Roy, 2013; Fanning & Cogger, 1998; Ozili, 2015) and majority of authors conclude that fraudulent activities include (1) change or manipulation of financial records, supporting documents or business transactions, (2) targeted manipulation or concealment of events, transactions, accounts and other information on the basis of which a company prepares its financial statements, (3) deliberate application of incorrect or misleading principles and methods used by a company to measure and report economic and commercial events, (4) targeted information concealment or targeted deception, and the presentation of misleading information, and (5) the use of aggressive accounting practices such as revenue timing.

Other authors studied relationship between market price, company earnings and book value of company shares, and due to convexity of function describing this relationship the results of this study show transfer of explanatory power from book value to earnings (Callao & Jarne, 2010).

Results show that companies that adopt IFRS tend to have higher earnings quality and lower cost of debt than those which do not adopt IFRS. These findings suggest that when unlisted companies issue bonds and borrow money, IFRS adoption contributes to decreasing the cost of this debt. International Financial Reporting Standards' implementation in the EU indeed covers two main objectives: to increase comparability of financial statements and parameters across the EU and to increase transparency of financial reporting (Chorvatovičová & Saxunová, 2016a).

Accounting frauds vary in their scope, context and positions of those who commit fraud in a company. Some types of fraud are specific to an industry or to certain groups of employees. Many frauds are centered on corporate senior management, but also frauds made by ordinary employees can be identified. Frauds committed by company employees include, for example, creation of fictitious costs and their subsequent reimbursement, creation of counterfeit employees to raise their own wages, creation of false suppliers and receiving payments for these suppliers for their own account and similar actions (Ozkul & Pamukc, 2012). If companies report higher net income (net earnings) but the following year or years the net income figures are lowered, this means that the net income would reverse. Earnings, that are the consequence of using selected accounting methods in the fiscal year, always reverse in the future. Therefore, earnings are of good quality if they do not reverse. This kind of manipulation is often referred to as earnings management (Chorvatovičová & Saxunová, 2016b).

Cressey's fraudulent scheme, so called fraudulent triangle, describes three major situational factors of fraudulent behaviors. Presence of these factors increases the likelihood of fraudulent behavior and these factors are: pressure, rationalization and opportunity (Cressey, 1973). Many contemporary authors have extended this model to a model of five factors, adding factors: knowledge and intent (Abbasi, Albrecht, Vance, & Hansen, 2012). The ambition to create more recent version of fraudulent triangle is presented by Rezaee, who introduces the concept of CRIME (C = cooking the books, R = recipes = instructions, I = incentives =

rewards / motivation, E = end results = results). Five different factors describe another, perhaps more modern, view of situational factors of fraudulent behavior (Rezaee, 2005).

Penman applies systematic approach to creation of fraud detection models in accounting. He presents main techniques and procedures to detect such an activity. It is important to focus on the quality of accounting and application of accounting procedures in a company (Penman, 2010). His work also focuses on steps how to distinguish apparent fraudulent behavior from accounting errors.

Apparent fraudulent activities also include, for example, overestimation of stock value, manipulation of capital expenditures (Messner, 2004), manipulation of company revenue (Roy, 2013), and earnings management (Tucker & Zarowin, 2006).

Although Apparao et al. (2009) and Albashrawi (2016) illustrate broad application of linear regression models in financial fraud detection, latest studies (Bhardwaj & Gupta, 2016; Kim, Baik, & Cho, 2016) point out to an increase in the use of data mining techniques in accounting fraud detection process. Performance comparison of these techniques compared to classical techniques (logistic regression, linear discriminative analysis) has never been performed, though.

Kirkos research (Kirkos, Spathis, & Manolopoulos, 2007) has proved that financial statements contain information which might help to detect accounting frauds. Accuracy of individual models is highest for Bayesian Belief network – 90.3%. These models contain 10 financial ratios calculated for 76 companies – out of which 38 companies have not reported any inaccuracies and 38 have reported inaccuracies. The accuracy rates of the Artificial Neural Network (ANN) model and the Decision Tree model are 80% and 73.6% respectively.

Lin (Lin, Chiu, Huang, & Yen, 2015) has applied machine learning method to 129 fraudulent and 447 non-fraudulent cases. Accuracy of testing dataset for decision trees is 90.3% and 92.80% for ANN. In the case of the Logistic Regression model the accuracy is 88.5%. Lin (Lin et al., 2015) research, though, has used 32 variables as input for the ANN model and 18 variables for Decision Trees. This is a high number of variables resulting in lower ease of application of these models to smaller businesses. Since some of these models include some hard-to-see data (e.g. number of internal auditors) or ambiguous qualitative features (e.g. kind of corporate ownership) it makes them more difficult to use in real life situations.

Most of data mining predictive models focus on fraudulent or intentional restatements for model building (e.g. Cecchini, Aytug, Koehler, & Pathak, 2010; Kim et al., 2016) and unintentional restatements have received very little attention in literature. Dutta et al. (2017) notice that there is no consensus on the best financial restatement model and its behavior depends purely on data structure and time. Dutta et al. research (2017) also uses data mining techniques as previously used by Li, but Dutta applies these techniques on unintentional restatement cases. Dutta's model (2017) has been tested for the period from 2001 to 2014 on 3404 non-restatement companies a 109 restatement companies and this model included 15 input variables. Its accuracy for decision trees is 69.3%, ANN accuracy is 64.4%, Support vector machine accuracy is 61.4%, Bayesian Belief network accuracy is 61.3% a Naïve Byes accuracy is only 44.8%.

Specific category of fraud detection research is represented by models focused on textual analysis of financial statements. These studies (Purda & Skillicorn, 2014; Y-J. Chen, Ch-H.

Wu, Y-M. Chen, H-Y. Li, & H-K. Chen, 2017) that are based on a pre-selected wordlist and frequency of use of these words in company financial statements, have achieved accuracy of 83–85% in prediction of fraudulent activities in company accounting. Authors of this study recommend using standard Beneish M-Score, F-Score, data mining techniques, and textual analysis as an addition to the word-list method.

Although differences among techniques are obvious, results are confirmed by their justification and use in various forms of fictitious frauds detection. Therefore, research in the field of financial restatements uses classical techniques along with modern, easily applicable data mining techniques (West, Bhattacharya, & Islam, 2014; I.P.-P.P.-M. Pervan, P. Pavić, & M. Pervan, 2014). These models use mainly data collected from financial statements for detection of accounting fraud (Gepp, 2015; Glancy & Yadav, 2011). Some models have also confirmed the existence of relationship between descriptive characteristics of a company, e.g. that the size of board of directors negatively impacts real activity-based earnings management (Kang & Kim, 2012).

The most common and studied model for accounting fraud detection is the Beneish model, also known as M-score (Beneish et al., 1999). Beneish model was been tested in several researches e. g. Ata and Seyrek (2009), Beneish, Lee, and Nichols (2012), Franceschetti and Koschtial (2013), Gepp (2015) or Herawati (2015). In the beginning of this research, the Beneish model behavior on a sample with financial restatement companies has been examined. Hence, it might be worth to test and determine prediction parameters of accuracy, sensitivity and specificity of existing fraud detection model on a sample of companies with financial restatements. Based on this testing it has been possible to determine to what extend prediction models developed for accounting frauds can be also used to detect accounting errors in financial reporting of a company.

1.1. Beneish M-Score

A model of eight factors is the M-score model presented by Beneish. This mathematical model uses 8 ratios to identify whether a company is manipulating its revenues or not. These ratios analyze data from financial statements, and after calculation of M-score it is possible to estimate the extent to which analyzed company manipulates its revenues (Beneish et al., 1999).

To build his model, between years 1987 until 1993, Beneish has studied 74 companies that manipulated their revenues and 2332 companies that did not manipulate their revenues.

Beneish model for detection of revenue and earnings manipulation uses following relationship:

$$\begin{aligned} M\text{-score} = & -4.84 + 0.92 * DSRI + 0.528 * GMI + \\ & 0.404 * AQI + 0.892 * SGI + 0.115 * DEPI - \\ & 0.172 * SGAI - 0.327 * LVGI + 4.679 * TATA. \end{aligned} \quad (1)$$

If the value of M-score is greater than -1.78 it can be assumed that the values in financial statements have been manipulated. Should the value of M-score be lower than -1.78 it is not expected that financial statements include fraud concerning revenue manipulation. The calculation of individual ratios is shown in the Table 1, where t stands for current period and $t-1$ for prior period.

Table 1. Beneish model variables – formulas

Variable	Index	Formula	Nb
Days Sales in Receivables Index	DSRI	$\frac{\frac{\text{Net receivables}_t}{\text{Sales}_t}}{\frac{\text{Net receivables}_{t-1}}{\text{Sales}_{t-1}}}$	(2)
Gross Margin Index	GMI	$\frac{\frac{\text{Sales}_t - \text{Cost of good sold}_t}{\text{Sales}_t}}{\frac{\text{Sales}_{t-1} - \text{Cost of good sold}_{t-1}}{\text{Sales}_{t-1}}}$	(3)
Asset Quality Index	AQI	$\frac{\frac{\text{Total assets}_t - \text{PPE}_t - \text{Current assets}_t - \text{Securities}_t}{\text{Total assets}_t}}{\frac{\text{Total assets}_{t-1} - \text{PPE}_{t-1} - \text{Current assets}_{t-1} - \text{Securities}_{t-1}}{\text{Total assets}_{t-1}}}$	(4)
Sales Growth Index	SQI	$\frac{\text{Sales}_t}{\text{Sales}_{t-1}}$	(5)
Depreciation Index	DEPI	$\frac{\frac{\frac{\text{Depreciation expense}_t}{\text{Depreciation expense}_t + \text{PPE}_t}}{\text{Depreciation expense}_{t-1}}}{\frac{\text{Depreciation expense}_{t-1} + \text{PPE}_{t-1}}$	(6)
Sales General and Administrative Expenses Index	SGAI	$\frac{\frac{\text{Selling, general and administrative expense}_t}{\text{Sales}_t}}{\frac{\text{Selling, general and administrative expense}_{t-1}}{\text{Sales}_{t-1}}}$	(7)
Leverage Index	LVGI	$\frac{\frac{\text{Total long term liabilities}_t + \text{Current liabilities}_t}{\text{Total assets}_t}}{\frac{\text{Total long term liabilities}_{t-1} + \text{Current liabilities}_{t-1}}{\text{Total assets}_{t-1}}}$	(8)
Total Accruals to Total Assets	TATA	$\frac{\text{Income from continuing operations}_t - \text{Cash flow from operations}_t}{\text{Total assets}_t}$	(9)

Note: Beneish et al. (1999).

Beneish' main assumption is that in the period prior to the manipulation period, manipulator companies report significantly higher growth than non-manipulator companies (medians are 29.4% v. 10.6%), suggesting that growth originates exogenously. Due to this reason the ratios in the Beneish model are ratios of two consecutive periods.

Other studies of the Beneish M-Score have focused on statistical significance testing of median variables entering this model (Ahmed & Naima, 2016; Franceschetti & Koschtial, 2013; Ramirez-Orellana, Martinez-Romero, & Mariño-Garrido, 2017; Repousis, 2016). All these studies focused only on financial fraud prediction and their results have confirmed the existence of statistical significance of following variables: Days Sales in Receivables Index

(DSRI), Asset Quality Index (AQI), Depreciation Index (DEPI), Selling, General and Administrative Expenses Index (SGAI), and would recommend reduction of Beneish variables only to these four variables. Another group of researchers (Paolone & Magazzino, 2014; Repousis, 2016; Özcan, 2018) have focused on the application of Beneish model in detecting fraudulent financial statements in specific countries such as Greece, Italy or Turkey. The last group of accounting detection studies consists of studies modifying basic Beneish M-score equation – by calculating new coefficients for its linear form, or by adjusting this equation to exponential form through logistic regression (Kara, Korpi, & Ugurlu, 2015). The main disadvantage of these studies is that even though adjusted models achieve high accuracy (81.1–86.63%), sensitivity (78.8–82.97%) and specificity (83.3–88.45%), these results have been obtained on training datasets, i.e. on data which have been used to develop these models.

Extending the study of Beneish (Beneish et al., 1999), Dechow, Ge, Larson, and Sloan (2011) classify companies with financial restatement and without financial restatement and develop F-score models using logit models. The F-score, in addition to the financial statements data, must be fed with various market data, off-balance sheet information and qualitative data about analyzed company.

Survey conducted on the Enron accounting fraud case by MacCarthy (2017) shows that Beneish M-Score compared to, for example, Altman's bankruptcy prediction model, could have successfully predicted Enron's fraudulent behavior between years 1997 and 2000. Research also shows that manipulation has been detected at majority of input variables feeding the Beneish model. Similar conclusion has been reached by Drábková (2015). Her research has confirmed, apart from M-score validity, also predictive strength of the CFEBT Model.

Variables of the Beneish model have become the basis for various alternative ways of detecting financial statements falsification, such as the Angus Z-score and other (Unegbu, 2013).

2. Research methodology

Research methodology along with more information about data samples and used statistical techniques is written in the following chapters.

2.1. Sample and data preparation

Financial information about 40 companies which have reported financial restatements (in a 10-K form) has been collected from the SEC database for U.S. based companies in the research sample and from annual reports of non-U. S based companies. Research period has been set to year of financial restatement publication and 10 previous years. For 19 companies in research sample, though, it has not been possible to collect complete financial reports over the past 10 years, so the sample contains 363 entries instead of 400 entries (EDGAR Online, 2018).

Input data have been collected for financial restatement year and ten years prior to this restatement publication. Collected data have been then divided into two groups: Dataset A and Dataset B.

The dataset A contains eight ratios relevant for calculation of the Beneish M-score. These ratios are:

- 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),
- Total Accruals to Total Assets (TATA).

The dataset B contains relative changes in variables listed below. Relative changes are calculated for prior year (before financial restatement) and two years back (after financial restatement). These variables are initial input variables to the Beneish M-score ratios included in Dataset A initial input variables and are as follows:

- Sales,
- Cost of goods sold,
- Selling, general and administrative expense,
- Income from continuing operations,
- Net income,
- Net receivables,
- Securities,
- Current assets,
- Property, plant and equipment,
- Total assets,
- Current liabilities,
- Total long-term liabilities,
- Depreciation expense,
- Cash flow from operating activities.

Calculation of relative change for prior year (2) is calculated by following formula:

$$Relative\ change_{t-1} = \frac{x_t - x_{t-1}}{x_{t-1}}. \quad (10)$$

Calculation of relative change dating two years backwards (3) is calculated by following formula:

$$Relative\ change_{t-2} = \frac{x_t - x_{t-2}}{x_{t-2}}, \quad (11)$$

where x_t is variable from Dataset A or Dataset B in current year; x_{t-1} is the same variable in prior year; x_{t-2} is two prior years variable.

Entire dataset of 40 companies is further divided into two groups: training dataset and testing dataset, with each containing 20 randomly chosen companies. Training dataset is used for development of logistic regression model and linear discriminant analysis model. Testing dataset is then used to test performance (accuracy, sensitivity and specificity) of developed models.

2.2. Logistic regression

Logistic regression estimates probability that a sample contains companies with financial restatements (if $y_k = 1$) and companies without financial restatements (if $y_k = 0$). Logistic regression works in the case of continuous independent variables. Logistic regression model (4) has following formula:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_1 x_{1,i} + \dots + \beta_k x_{k,i} + \dots + \beta_n x_{n,i})}}, \quad (12)$$

where $P(Y = 1|X)$ is probability that company financial statements include accounting error which will lead to financial restatement in the future; $x_{k,i}$ (whilst k varies from 1 to n) is an independent variable which represents variable from Dataset A or Dataset B; β coefficients represent estimated coefficients for independent variables.

Coefficients β are estimated by using maximum likelihood estimation for built-in function for generalized linear model in statistical software R. Fitted model is also described by Akaike Information Criterion and tested by Likelihood Ratio Statistic (LRS) (A. Field, J. Miles, & Z. Field, 2012).

2.3. Linear discriminant analysis

Linear discriminant analysis is a common method used to build predictive models. It has been previously also used by Altman to create his Z-score prediction model. Linear discriminant analysis is used to classify i companies which are described by n variables ($x_{1,i}, x_{2,i}, \dots, x_{n,i}$) into two groups: companies which would report financial restatements and companies which would not report financial restatements.

Model (5) for company i is defined in the following formula:

$$\text{score}_i = \alpha_1 x_{1,i} + \dots + \alpha_k x_{k,i} + \dots + \alpha_n x_{n,i}, \quad (13)$$

where variable $x_{k,i}$ (whilst k varies from 1 to n) represents variable k from Dataset A or Dataset B for company i .

Linear discriminant analysis estimates coefficients α used to calculate score for each company i . The estimation of coefficients α is done by a method based on maximization of variance between groups while minimizing the variance within groups. Based on calculated score which can be higher or lower than the threshold, company i is classified into one of two groups – companies which would report financial restatements and companies which would not report financial restatements (Field et al., 2012).

Final LDA model is tested on normality by Shapiro test, difference between groups is tested by ANOVA and MANOVA analysis and independence among variables by correlation analysis.

2.4. Description of model performance

Performance of each model – either the Beneish model or new own developed models – is evaluated by a number of parameters, such as: accuracy, sensitivity and specificity. Values of these parameters have been derived from the confusion matrix shown in Table 2. The table contains values of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) (Fawcett, 2006).

Values of variables used to determine performance of all predictive models are derived from the above matrix:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}; \quad (14)$$

$$Sensitivity = \frac{TP}{TP + FN}; \quad (15)$$

$$Specificity = \frac{TN}{FP + TN}. \quad (16)$$

Table 2. Confusion matrix

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Note: own source.

Accuracy (6) represents share of correctly predicted values relative to entire sample, Sensitivity (7) represents proportion of correctly identified companies with financial restatement, and Specificity (8) represents percentage of correctly identified companies without financial restatement. There are other indicators within performance model, but most authors agree that above parameters of Accuracy, Sensitivity and Specificity determine predictive performance of a model in the best way (Fawcett, 2006).

Above three performance indicators have been calculated for all developed models (logistic regression and LDA) as well as for both training dataset and testing datasets (Dataset A and Dataset B).

3. Results of analysis

Eight indicators of the Beneish model and relative changes for one and two prior years have been collected for 14 items from financial statements. To reduce the number of dimensions for tested models and to exclude some statistically insignificant variables, differences between companies with financial restatements and without restatements have been tested by double-sided t-test. Since no significant differences ($p\text{-value} < .05$) have been identified within groups with none of variables listed, tested models contained all variables from both datasets.

Descriptive analysis of both training dataset and testing dataset containing values including mean, median, and standard deviation, along with two-sample t test result, is shown in Table 3.

In his research, Beneish has identified that it is possible to determine 76% of manipulators accurately and 17.5% inaccurately as non-manipulators using only eight variables (MacCarthy, 2017) By applying the Beneish model on research sample of 40 companies, required predictive model performance as shown in Table 4 have not been achieved. Overall accuracy of the Beneish model is nearly 67% and is conditional on highly accurate data detection without financial restatements (75.63%). Out of total sample of 44 cases with financial restatements, only 3 cases of financial restatement have been correctly identified. This represents accuracy of the Beneish model less than 7%. Based on Table 4, it can be confirmed that performance of the Beneish model is approximately ten percent lower than estimated in the original Beneish research from 1999.

Table 3. Descriptive analysis for tested variables

Data-sets	Variables	No financial restatement			Financial restatement				
		Average	Median	Standard deviation	Average	Median	Standard deviation	P-value	Sign.
Dataset A	Days Sales in Receivables Index	1.113	0.978	1.219	0.951	0.987	0.256	0.835	
	Gross Margin Index	1.053	1.000	1.683	0.863	0.987	1.038	0.338	
	Asset Quality Index	1.566	1.008	2.537	1.096	1.041	0.412	0.470	
	Sales Growth Index	1.184	1.095	0.612	1.096	1.082	0.316	0.170	
	Depreciation Index	1.166	0.987	1.738	0.932	0.925	0.287	0.957	
	Sales General and Administrative Expenses Index	1.002	0.999	0.275	1.145	0.994	0.657	0.280	
	Leverage Index	1.068	1.001	0.522	1.026	0.976	0.351	0.486	
	Total Accruals to Total Assets	−0.032	−0.018	0.178	−0.073	−0.016	0.292	0.891	
Dataset B	Sales ($t-1$)	0.200	0.095	0.593	0.096	0.082	0.316	0.170	
	Cost of goods sold ($t-1$)	0.241	0.070	0.971	0.209	0.069	0.762	0.061	
	Selling, general and administrative expense ($t-1$)	0.184	0.101	0.395	0.142	0.104	0.324	0.954	
	Income from continuing operations ($t-1$)	0.189	0.105	4.067	−0.266	0.043	1.899	0.496	
	Net income ($t-1$)	0.096	0.102	4.998	9.930	0.101	63.122	0.426	
	Net receivables ($t-1$)	0.337	0.087	1.475	0.055	0.022	0.479	0.367	
	Securities ($t-1$)	0.795	0.000	6.314	0.100	0.000	0.755	0.610	
	Current assets ($t-1$)	0.233	0.090	1.124	0.103	0.060	0.304	0.700	
	Property, plant and equipment ($t-1$)	0.256	0.074	0.925	0.041	0.035	0.400	0.442	
	Total assets ($t-1$)	0.243	0.104	0.839	0.135	0.082	0.454	0.624	
	Current liabilities ($t-1$)	0.243	0.099	0.564	0.083	0.053	0.326	0.147	
	Total long term liabilities ($t-1$)	3.988	0.035	54.607	0.292	0.007	1.421	0.285	
	Depreciation ($t-1$)	0.308	0.124	1.459	0.347	0.121	0.999	0.392	
	Cash flow from operating activities ($t-1$)	0.289	0.068	8.700	0.094	0.032	1.593	0.741	
	Sales ($t-2$)	0.846	0.214	4.377	0.371	0.163	1.024	0.183	
	Cost of goods sold ($t-2$)	0.743	0.165	3.331	0.420	0.134	0.974	0.503	
	Selling, general and administrative expense ($t-2$)	0.543	0.260	1.784	0.526	0.219	1.420	0.750	
	Income from continuing operations ($t-2$)	0.498	0.206	5.087	−0.968	0.065	8.818	0.187	
	Net income ($t-2$)	0.609	0.258	5.708	24.455	0.090	168.529	0.300	
	Net receivables ($t-2$)	1.035	0.211	4.309	0.357	0.150	1.210	0.939	
	Securities ($t-2$)	1.070	0.000	7.363	1.315	0.000	5.209	0.771	
	Current assets ($t-2$)	0.560	0.185	2.244	0.299	0.156	0.507	0.669	
	Property, plant and equipment ($t-2$)	2.989	0.197	38.102	0.268	0.088	0.919	0.561	
	Total assets ($t-2$)	0.705	0.251	2.664	0.337	0.144	0.676	0.574	
	Current liabilities ($t-2$)	0.621	0.276	1.391	0.309	0.134	0.731	0.313	
	Total long term liabilities ($t-2$)	4.138	0.178	47.982	17.110	0.218	106.005	0.662	
	Depreciation ($t-2$)	1.718	0.266	17.425	0.657	0.315	1.830	0.641	
	Cash flow from operating activities ($t-2$)	−0.051	0.216	9.189	−3.447	0.081	17.111	0.553	

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Note: own calculation in MS Excel 2007 based on SEC data and company annual reports data.

Beneish M-Score is a method which can be used to detect companies with tendency to commit fraud within their financial statements (Beneish et al., 2012). Unlike accounting fraud, restatement of company financial statements is a revision and correction of one or more company's previous financial statements. These financial restatements are a result of previous accounting error, non-compliance with GAAP or simple clerical error.

Ratios used in Dataset A do not have significant influence on financial restatement detection in neither logistic regression nor in linear discriminant analysis. Probability for each ratio is not lower than 0.05 which represents significance level. Residual standard error for logistic regression model is 118.7 on 167 degrees of freedom and Akaike Information Criterion is 134.7. Results from likelihood ratio statistic show that this logistic model is not more efficient than sample mean ($\chi^2(1) = 5.6833$, $p = .5772$).

ANOVA and MANOVA results for linear discriminant analysis show no significant differences between groups with financial restatements and without financial restatements on group level ($W = 0.9291$, $p = .9715$) as well as on variable level, as shown in Table 2. Correlation ($|x| > 0.5$) among Dataset A ratios is not identified in the tested model.

Table 4. Descriptive analysis for tested variables for Beneish model

		Actual		Beneish model performance		
		Restatements	Non-Restatements	Accuracy	Sensitivity	Specificity
Prediction	Restatements	3	77	66.67%	6.38%	75.63%
	Non-Restatements	44	239			

Note: own calculation in MS Excel 2007 based on SEC data and company annual reports data.

Table 5. Results of logistic regression and linear discriminant analysis for dataset A

Variable	Logistic regression				Linear Discriminant Analysis			
	Coeff.	Z value	Pr(> z)	Sign.	Coeff.	F value	Pr(>F)	Sign.
Days Sales in Receivables Index	-0.7187	-0.839	0.401		0.8793	0.181	0.6724	
Gross Margin Index	-0.2288	-0.968	0.333		-6.5644	0.384	0.5382	
Asset Quality Index	-0.3329	-0.695	0.487		3.2477	0.718	0.4010	
Sales Growth Index	-0.9871	-0.967	0.333		7.1669	0.536	0.4676	
Depreciation Index	0.1104	0.188	0.851		-1.9679	0.087	0.7693	
Sales General and Administrative Expenses Index	0.5863	0.792	0.429		8.3983	0.401	0.5295	
Leverage Index	-0.3042	-0.259	0.796		1.6815	0.088	0.7673	
Total Accruals to Total Assets	0.4455	0.227	0.820		-18.3560	0.915	0.3436	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Note: own calculation in R studio based on SEC data and company annual reports data.

Results in Table 5 show that ratios used in Dataset A are not effective for detection of financial restatements as much as they are for detection of accounting frauds. Financial restatements are necessary to be reported when previous period financial statements include

any inaccuracy. In such cases, differences in entries from financial statements can be more effective to be analyzed rather than sophisticated financial ratios as used in Beneish model.

Results of logistic regression model from Table 6 indicate that selling, general and administrative expenses ($z = -2.528$, $p < .05$), income from continuing operations ($z = -1.75$, $p < .1$), net income ($z = 2.167$, $p < .05$) and current liabilities ($z = -2.982$, $p < .01$) all with relative change (current period to previous period) have significant effect on detection of companies with financial restatements. If variable's coefficient is negative, likelihood of financial restatement is lower. Residual standard error for logistic regression model is 197.43 on 171 degrees of freedom, Akaike Information Criterion is 205.43. Furthermore, results from this calculation show that this logistic model is efficient ($\chi^2(1) = 73.048$, $p < .005$).

Results of LDA model from Dataset B variables are represented in Table 7. Results of this model indicate that none from tested variables have significant effect on detection of companies with financial restatements. On variable level differences between groups with financial restatements and without financial restatements are not significant.

Table 6. Results of logistic regression for dataset B

Variable	Coeff.	Z value	Pr(> z)	Sign.
Selling, general and administrative expense ($t-1$)	-2.1482	-2.528	0.01147	*
Income from continuing operations ($t-1$)	-0.2777	-1.750	0.08007	.
Net income ($t-1$)	0.2804	2.167	0.03021	*
Current liabilities ($t-1$)	-2.2513	-2.982	0.00286	**

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Note: own calculation in R studio based on U.S. Securities and Exchange Commission Companies and from annual reports.

Table 7. Results of linear discriminant analysis for dataset B

Variable	Coeff.	F value	Pr(>F)	Sign.
Cost of goods sold ($t-1$)	-27.5208	1.2126	0.28050	
Current liabilities ($t-1$)	-5.1724	1.8950	0.17990	
Depreciation ($t-1$)	29.2749	0.6674	0.42110	
Income from continuing operations ($t-1$)	6.0012	1.2429	0.27470	
Cost of goods sold ($t-2$)	3.9989	0.2766	0.60320	
Selling, general and administrative expense ($t-2$)	7.9949	0.0047	0.94610	
Property, plant and equipment ($t-2$)	-8.7047	0.2958	0.59100	
Depreciation ($t-2$)	-7.1253	0.6102	0.44150	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Note: own calculation in R studio based on U.S. Securities and Exchange Commission Companies and from annual reports.

MANOVA results for LDA model, though, show significant differences between groups with financial restatements and without financial restatements ($W = 0.2242$, $p < .05$). Figure 1 shows how combination of variables classified training dataset into two groups. Like for the

LDA model based on Dataset A, there is also no correlation ($|x| > 0.5$) among Dataset B in tested model. Due to this reason, tested model is accepted as significant for unintentional financial restatements detection.

Models are established on training data and their performance attributes such as accuracy, sensitivity and specificity which are shown in Table 8. These performance attributes are then recalculated on testing dataset and after this calculation these performance attributes worsen. In both cases (for logistic regression and LDA) and both data (training and testing) models based on Dataset B have higher accuracy and specificity than models based on Dataset A. It indicates that Dataset A (the Beneish ratios) are not very accurate for prediction of unintentional financial restatements in companies.

Cut-off level for logistic regression and LDA is determined as maximum achieved value of sum of sensitivity and specificity. For logistic regression and Dataset A, this value equals 0.127 and for LDA it equals 14.038. Should the output value of a model be higher than these two values, it can be assumed that analyzes financial statement includes an unintentional error that requires financial restatement. Similarly, for Dataset B, logistic regression has been determined to be 0.555 and LDA to be 1.351.

Training data for Dataset B achieve better results for LDA model. On the other side, testing data show logistic regression model is more stable because accuracy decreased by only

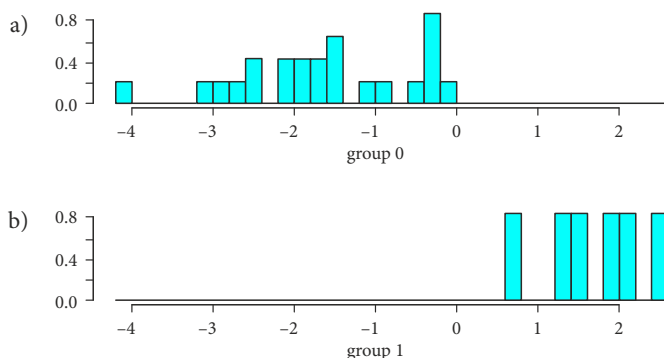


Figure 1. Density plot with both a) “without financial restatement” (group 0) and b) “with financial restatement” (group 1)

Table 8. Comparison of performance for logistic regression and linear discriminant analysis models

Variables	Data	Logistic regression			Linear discriminant analysis		
		Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Dataset A	Training	64.57%	65.00%	64.51%	82.35%	0.00%	97.67%
	Testing	61.93%	33.33%	67.17%	57.78%	41.67%	61.26%
Dataset B	Training	81.14%	45.00%	85.81%	100.00%	100.00%	100.00%
	Testing	70.96%	25.00%	79.38%	62.22%	41.66%	66.67%

Note: own calculation in R studio.

approx. 10% and specificity by only approx. 6%. Logistic regression model is very stable with Dataset A variables and its accuracy decreased by not more than 4%. It shows that logistic regression model has higher potential for prediction of unintentional financial restatements in companies. Logistic regression model has achieved better performance in all performance parameters than original Beneish model. Accuracy of original Beneish model has been approx. 67%, sensitivity 6.38% and specificity 75.63%.

It can be determined that relevant models for financial restatement detection are:

- LDA model for Dataset B for variables: Cost of goods sold for $t-1$, Current liabilities for $t-1$, Depreciation for $t-1$, Income from continuing operations for $t-1$, Cost of goods sold for $t-2$, Selling, general and administrative expenses for $t-2$, Property, plant and equipment for $t-2$ and Depreciation for $t-2$,
- Logistic regression model for Dataset B only for significant variables: Selling, general and administrative expense for $t-1$, Income from continuing operations for $t-1$, Net income for $t-1$ and Current liabilities for $t-1$.

For easier application and interpretation of these models, it is meaningful to rewrite them in a form of equation. LDA model for Dataset B has following form (17):

$$\begin{aligned} LDA \text{ Score} = & -27.52 * \text{Cost of goods sold}_{t-1} - 5.17 * \text{Current liabilities}_{t-1} + \\ & 29.27 * \text{Depreciation}_{t-1} + 6 * \text{Income continuing operations}_{t-1} + 4 * \\ & \text{Cost of goods sold}_{t-2} + 7.99 * \text{Selling, general and administrative expense}_{t-2} - \\ & 8.7 * \text{Property, plant and equipment}_{t-2} - 7.13 * \text{Depreciation}_{t-2}. \end{aligned} \quad (17)$$

If LDA Score result is greater than 1.351, there is increased probability of an accounting error that might lead to financial restatement in analyzed company. If LDA Score result is lower than 1.351, there is no assumption of an accounting error leading to financial restatement.

Similarly, for logistic regression model for significant variables applied on Dataset B, an equation that has following form (18) has been created:

$$LR \text{ Score} = \frac{1}{1 + e^{-(2.15 * x_1 - 0.28 * x_2 + 0.28 * x_3 - 2.25 * x_4)}}. \quad (18)$$

Where variable x_1 are Selling, general and administrative expenses for $t-1$, x_2 is Income from continuing operations for $t-1$, x_3 is Net income for $t-1$ a x_4 are Current liabilities for $t-1$. If LR score is greater than 0.555, there is likelihood that company's financial reporting includes accounting errors that have not been discovered yet and thus there is probability that the company will have to prepare financial restatement in the future. If LR Score is lower than 0.555, this probability does not exist or is very low.

Such models, either newly developed own models for financial restatement prediction or Beneish or Altman models, can be easily applied and quickly interpreted. These attributes help users to apply new predictive models to business practice. In academic field, these equations might allow users to easily re-test these models and analyze newly developed models for possible corrections or updates.

4. Discussion

This manuscript identifies several determinants that cause companies to restate previously published financial statements due to discovery of unintentional accounting errors. For newly developed LDA model, these determinants have been identified as *Cost of goods sold*_{*t*-1}, *Current liabilities*_{*t*-1}, *Depreciation*_{*t*-1}, *Income from continuing operations*_{*t*-1}, *Cost of goods sold*_{*t*-2}, *Selling, general and administrative expenses*_{*t*-2}, *Property, plant and equipment*_{*t*-2} and *Depreciation*_{*t*-2}. In case of logistic regression model, determinants that have been identified are *Selling, general and administrative expenses*_{*t*-1}, *Income from continuing operations*_{*t*-1}, *Net income*_{*t*-1} and *Current liabilities*_{*t*-1}. Research conducted in this manuscript focuses on debt-related accounts, cash accounts and revenue recognition accounts which is in line with previously conducted studies.

This study has proved that it is possible to develop financial restatements prediction models based on information from publicly available financial statements of companies. Intentional manipulation of financial statements, as well as unintentional errors in company financial reporting, are possible to be uncovered by accompanying mathematical phenomenon. This mathematical phenomenon can be predicted based on analysis of debt-related accounts, cash accounts and revenue recognition accounts.

Moreover, as this research indicates, utilization of relatively small sample of companies (40 restatement companies used in this research) in combination with classical mathematical and statistical techniques – of logistic regression and linear discriminant analysis – provide comparable, if not better results than studies that use higher number of sample companies in sample or mathematical and statistical techniques based on machine learning. Based on results from this manuscript and corresponding with observations of Beneish, it is determined that it is possible to develop prediction models with similar or even better accuracy by using lower number of research companies by utilization of classical mathematical techniques.

Models developed in this research, like the Beneish M score and the Altman Z score, are easily applicable on available data. In line with these two fundamental studies, also this manuscript provides newly developed models in a form of equation that can be easily tested – by both academics and business professionals. Developed models can be widely used by both internal and external users of financial statements, who would like to determine if financial statements of analysed company include accounting errors or not, thanks to easily interpretable results in equation form.

Possible limitation of newly created models in this research is the fact that that these models use publicly available information as input variables, and therefore cannot be applied to private companies whose financial statements are not publicly available to pull input data from. Companies included in the research sample are also from different time periods over the course of last 20 years, which could have distorted results of this study to a limited extend. The ability to focus on same time interval, e.g. horizon of the last ten years, can contribute to greater relevance of developed models and, therefore, future studies might focus also on this issue.

Since new models have been tested mainly on a sample of US-listed companies and thus mainly in line with US GAAP, it is questionable to what extent these models could be also

applied to companies from other regions that are preparing their financial statements under other accounting standards (e.g. IFRS, etc.). Application of new models on small and medium sized companies, that prepare their financial statements in line with local GAAP, is questionable and should be analyzed and tested in further studies.

As mentioned in Introduction of this manuscript, there are 4 major financial restatement areas: debt-related accounts, cash accounts, tax expense accounts and revenue recognition accounts. Models developed in this study exclude tax expense accounts as these accounts have been proved to be statistically insignificant. Hence prediction possibilities to uncover accounting errors in the field of tax expense leading up to financial restatements in the future are not considered in this research. Authors of this research also recommend tax accounts to be considered in future studies in the area of financial restatements and their determinants.

The main contribution of this research is that it is one of pioneering studies conducted in financial restatements area and subsequent prediction model development. Even though financial restatements pose negative impact on perception of a company in both public and investor eye and have direct impact on company market value, so far there has been only one study that focuses on prediction of financial restatements. This research fills an academic research gap of financial restatement phenomenon and provides suggestions for future studies in financial restatement prediction.

Conclusions

Financial restatements are not studied in much detail in academic literature. Despite high impact of financial restatements on trust of company stakeholders and investors, almost no authors address the topic from the point of possible detection and prediction. Present studies focus mainly on detection of accounting frauds through data mining techniques and logistic regression. Outcomes of these studies are not easily applicable on real-life situations as they focus on describing possible ways how to build up prediction models for accounting fraud detection. Such an approach is not very effective in fraud detection process since development and application of such models is time consuming and expensive.

Creation of mathematical models for investigation of statistical dependency mainly uses methods of discriminant analysis and logistic regression. Results of models developed in this research indicate that despite significant randomness it is possible to predict financial restatements with relatively high accuracy. Analysis performed in this study has proved that development of more accurate models for financial restatement prediction may be achieved when using year-over-year changes in financial data variables as input variables, as opposed to using sophisticated financial ratios as in the Beneish model. Models based on year-over-year changes in financial data variables are more accurate by approximately 20 percent on training dataset and by 5 to 10 percent on testing dataset. Logistic regression model which has been developed for year-over-year changes in financial data variables also achieves predictive precision comparable with the original Beneish model, and even surpasses its accuracy when tested on research sample of this study. Based on this fact it can be concluded that the Beneish model can be considered as not precise enough for unintentional restatements

detection. The need of further research that would create new prediction models in the area of financial restatements cannot be pointed out more.

New model created by linear discriminatory analysis has proved that reliance on modeling only p-values cannot lead to results. There have been no significant differences between the two groups of training datasets, however, final combination of variables has resulted in high differentiation between the two groups.

Among new developed models it has been noticed that higher accuracy of training dataset has been established in models created by linear discriminant analysis. Testing has proved, though, that higher stability and accuracy have been achieved by using logic regression models (from 5 up to 10 percent higher).

New developed models of this research would surely require further testing and analysis as there are not many studies on this topic. This manuscript is first of a kind to have used linear discriminant analysis and first to have tested the Beneish model on a sample of financial restatement companies.

Prediction of unintentional financial restatements in financial reporting is a very complex process that involves significant data collection. Once a financial restatement is prepared and published, the challenge of obtaining original pre-restatement data proves to be difficult and time consuming. Due to this reason, it is essential to continue in this research by collecting further data – especially with focus on collecting data per specific sectors, and creation of models for these sectors. Each sector has its own specifications and their importance has been already proved by the creation of bankruptcy prediction models separately for each business sector.

Possible limitation of models developed in this manuscript is the fact that these models use publicly available information of mainly US-listed companies as input variables. These input variables represent different time periods over the course of the last 20 years. During this period, a lot of changes in legislation have been introduced. Moreover, companies included in this research sample are preparing their financial statements in line with US GAAP. Hence, application of prediction models developed in this research on companies that are preparing their financial statements under other accounting standards (e.g. IFRS) is questionable and should be further tested. Last area of limitation is the fact that tax accounts, accounts from major financial restatements area, are not considered in this research. In a case when financial restatement is to be prepared because of an error in tax accounts, models developed in this study are not able to detect them.

There is currently low number of expert studies to identify unintentional financial restatements determinants, whilst markets responses to these reporting changes in statements in a significant manner. Due to this reason, it is necessary to continue to address this issue and not to brand classical mathematical techniques as outdated, since their interpretation and application is much easier than application of advanced data mining techniques. Accuracy of these rather traditional techniques even surpasses some studies that have used data mining techniques.

Disclosure statement

The authors (Lenka Papíková and Mário Papík) have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

- Abbasi, A., Albrecht, C., Vance, A., & Hansen J. (2012). Metafraud: A meta-learning framework for detecting financial fraud. *MIS Quart Manage Inf Syst MIS Quarterly: Management Information Systems*, 36(4), 1293-1327. <https://doi.org/10.2307/41703508>
- Ahmed, T., & Naima, J. (2016). Detection and analysis of probable earnings manipulation by firms in a developing country. *Expert Systems with Application Asian Journal of Business and Accounting*, 9(1), 59-82.
- Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: a decade review from 2004 to 2015. *Journal of Data Science*, 14(3), 553-570.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23(4), 589-609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Apparao, G., Singh, A., Rao, G. S., Bhavani, B. L., Eswar, K., & Rajani, D. (2009). Financial statement fraud detection by data mining. *Corporate governance*, 3(1), 159-163. Retrieved from <http://www.ijana.in/papers/3.3.pdf>
- Ata, A. & Seyrek, I. (2009). The use of data mining techniques in detecting fraudulent financial statements: an application on manufacturing firms. *The Journal of Faculty of Economics and Administrative Sciences*, 14(2), 157-170.
- Audit Analytics. (2015). *2014 financial restatements: a fourteen year comparison*. USA. Sutton: Audit Analytics.
- Bateineh, H., Abuaddous, M., & Alabood, E. (2018). The effect of family ownership and board characteristics on earnings management: evidence from Jordan. *Academy of Accounting and Financial Studies Journal*, 22(4), 1-17.
- Beneish, M. D., Lee, C., Press, E., Whaley, B., Zmijewski, M., & Cisilino, P. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24-36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Beneish, M. D., Lee, C. M. C., & Nichols, D. C. (2012). Fraud detection and expected return. *Financial Analysts Journal*, 69(5), 14-14. <https://doi.org/10.2139/ssrn.1998387>
- Bhardwaj, A., & Gupta, R. (2016). Financial frauds: data mining based detection-a comprehensive survey. *International Journal of Computer Applications*, 156. <https://doi.org/10.5120/ijca2016912538>
- Bhasin, M. L. (2013). Corporate governance and forensic accountants' role: global regulatory action scenario. *International Journal of Accounting Research*, 1(1), 1-19. <https://doi.org/10.2139/ssrn.2676468>
- Callao, S., & Jarne, J. (2010). Have IFRS affected earnings management in the European Union. *Accounting in Europe*, 7(2), 159-189. <https://doi.org/10.1080/17449480.2010.511896>
- Cressey, D. R. (1973). *Other people's money: study in the social psychology of embezzlement*. Montclair, N. J., Patterson Smith.
- Cecchini, M., Aytug H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science*, 56(7), 1146-1160. <https://doi.org/10.1287/mnsc.1100.1174>

- Chen, K. Y., & Zhou, J. (2010). Audit committee, board characteristics, and auditor switch decisions by Andersen's clients. *Contemporary Accounting Research*, 24(4), 1,085-1,117.
<https://doi.org/10.1506/car.24.4.2>
- Chen, Y.-J., Wu, Ch.-H., Chen, Y.-M., Li, H.-Y., & Chen, H.-K. (2017). Enhancement of fraud detection for narratives in annual reports. *International Journal of Accounting Information Systems*, 26, 32-45.
<https://doi.org/10.1016/j.accinf.2017.06.004>
- Chorvatovičová, L., & Saxunová, D. (2016a). Earnings management after IFRS implementations across the European Union. In *Bilaterálne akademické fórum – Slovensko a Francúzsko v perspektíve zahraničnej politiky SR* (pp. 84-90). Czech Republic: Wolters Kluwer. Retrieved from https://www.fm.uniba.sk/fileadmin/fm/Veda/Forum_bilateralne/Zbornik_Bilateralne_akademicke_forum_Slovensko_a_Francuzsko_v_perspektive_zahranicnej_politiky_SR_FMUK_2016_finale.pdf
- Chorvatovičová, L., & Saxunová, D. (2016b). Usefulness of financial statements and annual reports in the process of accounting fraud detection. In *Managing Global Changes: Proceedings of the joint international conference* (pp. 233-247). Croatia: University of Primorska Press. Retrieved from <http://www.hippocampus.si/ISBN/978-961-6984-81-2/130.pdf>
- Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements*: predicting material accounting misstatements. *Contemporary Accounting Research*, 28, 17-82. <https://doi.org/10.1111/j.1911-3846.2010.01041.x>
- Drábková, Z. (2015). Analysis of possibilities of detecting the manipulation of financial statements in terms of the IFRS and Czech accounting standards. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63, 1859-1866. <https://doi.org/10.11118/actaun201563061859>
- Duska, R., Duska, B., & Ragatz, A. (2011). *Accounting ethics*. New Jersey, USA: John Wiley & Sons Publishing. <https://doi.org/10.1002/9781444395907>
- Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems with Applications*, 90, 374-393. <https://doi.org/10.1016/j.eswa.2017.08.030>
- EDGAR Online. (2018). *List of companies*. 10-K.
- Fanning, K., & Cogger, K. (1998). Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance & Management*. [https://doi.org/10.1002/\(SICI\)1099-1174\(199803\)7:1<21::AID-ISAF138>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1174(199803)7:1<21::AID-ISAF138>3.0.CO;2-K)
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861-874.
<https://doi.org/10.1016/j.patrec.2005.10.010>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. United Kingdom. London: SAGE Publications Ltd.
- Franceschetti, B. M., & Koschtial, C. (2013). Do bankrupt companies manipulate earnings more than the non-bankrupt ones? *Journal of Finance and Accountancy*, 12, 4-24.
- Gepp, A. (2015). *Financial statement fraud detection using supervised learning methods*. Retrieved from <http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1227&context=theses>
- Girgenti, R., & Hedley, T. (2011). *Managing the risk of fraud and misconduct*. New York, USA: McGraw Hill.
- Glancy, F. H., & Yadav S. B. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems*, 50, 595-601. <https://doi.org/10.1016/j.dss.2010.08.010>
- Gleason, C. A., Jenkins, N. T., & Johnson, W. B. (2008). The contagion effects of accounting restatements. *The Accounting Review*, 83(1), 83-110. <https://doi.org/10.2308/accr.2008.83.1.83>
- Herawati, T. N. (2015). Application of Beneish M-Score models and data mining to detect financial fraud. *Procedia – Social and Behavioral Sciences*, 211, 924-930.
<https://doi.org/10.1016/j.sbspro.2015.11.122>
- La Torre, I. (2009). *Creative accounting exposed*. New York, USA: Palgrave Macmillan.

- Lemus, E. (2014). The financial collapse of the Enron corporation and its impact in the United States Capital Market. *Global Journal of Management and Business Research*, 14(4), 40-51. Retrieved from <https://journalofbusiness.org/index.php/GJMBR/article/view/1539/1442>
- Lin, C.-C., Chiu, A.-A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 89, 459-470. <https://doi.org/10.1016/j.knosys.2015.08.011>
- Kang, S.-A., & Kim, Y.-S. (2012). Effect of corporate governance on real activity-based earnings management: evidence from Korea. *Journal of Business Economics and Management*, 13, 29-52. <https://doi.org/10.3846/16111699.2011.620164>
- Kara, E., Korpi, M., & Ugurlu, M. (2015). Using Beneish model in identifying accounting manipulation: an empirical study in BIST manufacturing industry sector. *Journal of Accounting, Finance and Auditing Studies*, 1(1), 21-39.
- Kim, Y. J., Baik, B., & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems with Applications*, 62, 32-43. <https://doi.org/10.1016/j.eswa.2016.06.016>
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32, 995-1003. <https://doi.org/10.1016/j.eswa.2006.02.016>
- MacCarthy, J. (2017). Using Altman Z-score and Beneish M-score models to detect financial fraud and corporate failure: a case study of Enron corporation. *International Journal of Finance and Accounting*, 6, 159-166. <https://doi.org/10.5923/j.ijfa.20170606.01>
- Messner, M. (2004). Exploring options in forensic accounting. *National Public Accountants*, 19-20.
- Nakashima, M. (2017). Can The Fraud Triangle predict accounting fraud?: Evidence from Japan. *Proceedings: The 8th International Conference of the Japanese Accounting Review*. Japan: Rokkoudai Campus of Kobe University. Retrieved from http://www.rieb.kobe-u.ac.jp/tjar/conference/8th/CC2_MasumiNAKASHIMA.pdf
- Oyebisi, O., Wisdom, O., Olusogo, O., & Ifeoluwa, O. (2018). Forensic accounting and fraud prevention and detection in Nigerian banking industry. *COJ Reviews & Research*, 1(1), 1-8. <http://dx.doi.org/10.31031/cojrr.2018.01.000504>
- Özcan, A. (2018). The use of Beneish model in forensic accounting: evidence from Turkey. *Journal of Applied Economics and Business Research*, 8(1), 57-67.
- Ozili, P. K. (2015). Forensic accounting and fraud: a review of literature and policy implications. *International Journal of Accounting and Economic Studies*, 3(1), 63-68. <http://doi.org/10.2139/ssrn.2628554>
- Ozkul, F. U., & Pamukcu, A. (2012). Fraud detection and forensic accounting. *Emerging Fraud*, 19-41. https://doi.org/10.1007/978-3-642-20826-3_2
- Paolone, F., & Magazzino, C. (2014). Earnings manipulation among the main industrial sectors. evidence from Italy. *Business and Management Sciences International Quarterly Review*, 5(4), 253-261.
- Penman, S. H. (2010). *Financial statement analysis and security valuation*. New York, USA: McGraw Hill.
- Pervan, I., Pavić, P., & Pervan, M. (2014). Firm financial distress prediction with statistical methods: prediction accuracy improvements based on the financial data restatements. *8th International Days of Statistics and Economics*. Retrieved from https://msed.vse.cz/msed_2014/article/278-Pervan-Ivica-paper.pdf
- Purda, L., & Skillicorn, D. (2014). Accounting variables, deception, and a bag of words: assessing the tools of fraud detection. *Contemporary Accounting Research*, 32(3). <https://doi.org/10.1111/1911-3846.12089>

- Ramirez-Orellana, A., Martinez-Romero, M., & Mariño-Garrido, T. (2017). Measuring fraud and earnings management by a case of study: Evidence from an international family business. *European Journal of Family Business*, 7(1-2), 41-53. <https://doi.org/10.1016/j.ejfb.2017.10.001>
- Rezaee, Z. (2005). Causes, consequences and deterrence of financial statement fraud. *Critical Perspectives on Accounting*, 16(3), 277-298. [https://doi.org/10.1016/S1045-2354\(03\)00072-8](https://doi.org/10.1016/S1045-2354(03)00072-8)
- Repousis, S. (2016). Using Beneish model to detect corporate financial statement fraud in Greece. *Journal of Financial Crime*, 23(4), 1063-1073. <https://doi.org/10.1108/JFC-11-2014-0055>
- Roy, M. (2013). *Financial statement fraud detection and prevention*. Amsterdam, Netherlands: World Press.
- Sadaf, R., Oláh, J., Popp, J., & Mate, D. (2018). an investigation of the influence of the worldwide governance and competitiveness on accounting fraud cases: a cross-country perspective. *Sustainability*, 10(3), 1-11. <https://doi.org/10.3390/su10030588>
- Savčuk, O. (2007). Internal audit efficiency evaluation principles. *Journal of Business Economics and Management*, 8, 275-284. <https://doi.org/10.3846/16111699.2007.9636180>
- Song, M., Oshiro, N., Shuto, A. (2016). Predicting accounting fraud: evidence from Japan. *The Japanese Accounting Review*, 6, 17-63. <https://doi.org/10.11640/tjar.6.2016.01>
- Stanley, J. D., & DeZoort, F. T. (2007). Audit firm tenure and financial restatements: an analysis of industry specialization and fee effects. *Journal of Accounting and Public Policy*, 26(2), 131-159. <https://doi.org/10.1016/j.jaccpubpol.2007.02.003>
- Tucker, J. W., & Zarowin, P. A. (2006). Does income smoothing improve earnings informativeness? *The Accounting Review*, 81(1), 251-270. <https://doi.org/10.2308/accr.2006.81.1.251>
- Unegbu, A. O. (2013). Advances in modeling for falsified financial statements. *International Journal of Finance and Accounting*, 2, 37-54. Retrieved from <http://article.sapub.org/10.5923.j.ijfa.20130201.06.html#Sec1>
- West, J., Bhattacharya, M., & Islam, R. (2014). Intelligent financial fraud detection practices: an investigation. In *International Conference on Security and Privacy in Communication Systems* (pp. 186-203). Springer. <https://doi.org/10.1016/j.cose.2015.09.005>