International Journal of Forecasting 39 (2023) 123–143

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Influence of earnings management on forecasting corporate failure

David Veganzones [a,*], Eric Séverin [b], Souhir Chlibi [b]

[a] Inseec U Research Center, ESCE International Business School, Paris, France
[b] IAE LILLE, Université de Lille, ULR 4999 LUMEN, Lille, France

ARTICLE INFO

Keywords:
Forecasting
Corporate failure
Earnings management
Threshold model
Finance

ABSTRACT

This paper studies the relationship between corporate failure forecasting and earnings management variables. Using a new threshold model approach that separates samples into different regimes according to a threshold variable, the authors examine regimes to evaluate the prediction capacities of earnings management variables. By proposing this threshold model and applying it innovatively, this research reveals boundaries within which earnings management variables can yield superior corporate failure forecasting. The inclusion of earnings management variables in corporate failure models improves failure prediction capacities for firms that manipulate substantial earnings. Furthermore, an accruals-based variable improves predictions of failed firms, but the real activities-based variable improves predictions of non-failed firms. These findings highlight the importance of indicators of the magnitude of earnings management and the tools used to improve the performance of corporate failure models. The proposed model can determine the predictive power of particular explanatory variables to forecast corporate failure.

1. Introduction

Topics of corporate failure prediction and earnings management have drawn substantial attention in forecasting, accounting, and financing domains. The ability to predict which firms will survive and thus which will meet their obligations has become increasingly important in recent decades owing to an increasing number of corporate failures. Failure prediction is key to the daily activities of many financial institutions because it is essential to the sound functioning of financial systems (Forgione & Migliardo, 2018). Similarly, earnings management evokes academic and practical interest due to the risks associated with intentional efforts to alter financial statements to signal desired, rather than real information, as when firms distort financial information, fail to report information transparently, or mislead shareholders and outsiders (García Lara, Garcia Osma, & Mora, 2005). Major accounting scandals at companies such as Enron, Parmalat, and Bankia have negatively affected accounting and financial institutions and eroded public confidence. In research that investigates firms' financial situations in conjunction with earnings management practices, evidence indicates that when firms experience financial distress that eventually leads to failure, their managers have incentives to conduct financial manipulation and avoid reporting true financial conditions. Notably, earnings management is pervasive more among bankrupt firms than among healthy entities (Dutzi & Rausch, 2016; Hassanpour & Ardakani, 2017).

Information about earnings manipulation thus provides a way to distinguish sound firms from unsound ones. Because financial accounts can reveal the causes of most corporate failures, most corporate failure models have relied solely on financial ratios calculated from firms' financial statements. It is not surprising that financial information sources dominate corporate failure prediction because the relationship between financial information and corporate failure is irrefutable and offers

* Corresponding author.
E-mail address: dveganzones@esce.fr (D. Veganzones).

https://doi.org/10.1016/j.ijforecast.2021.09.006

objective measures (Micha, 1984). However, authors have contested the idea that a model can predict failure using evidence only from financial ratios because certain indicators of corporate failure are not reflected in those ratios (du Jardin, 2009). Non-financial variables, such as firms' strategies, corporate governance, and macroeconomic indicators, which are not embodied in accounting data can exert significant influence in predicting the risk of failure. In turn, various authors have investigated the usefulness of models that, while based on financial ratios, also include complementary non-financial information (Brédart, Séverin, & Veganzones, 2021; Ciampi, 2015; Mare, 2015). Such non-financial variables, particularly those related to firms' characteristics, are difficult to gather though and can rather be inefficient, such that the marginal gain in accuracy obtained by inclusion is minimal compared with the marginal cost of collecting the information. Accordingly, we argue that corporate failure forecasting literature could be extended by integrating a novel class of predictors, i.e., earnings management variables that reflect the presence of accounting anomalies in annual accounts, such that the combined use of financial ratios and earnings management variables produces a model that incorporates dual financial analysis. Financial ratios indicate firms' financial dimensions, and the earnings management information reflects the quality of the informational financial content. Therefore, the earnings management information may be an ideal complement to traditional corporate failure models. Moreover, because earnings management information can be calculated for any type of firm that publishes its financial accounts, we suggest applying the earnings-management-based models easily to real-world scenarios. However, because the predictive power of earnings management information cannot be assumed automatically, it is limited by regulatory scrutiny of accounts, accounting flexibility of firms, and cost (Gunny, 2010). Thus, it becomes essential to evaluate the extent to which the use of earnings management information can improve corporate failure models.

Therefore, in this study, we investigate the usefulness of earnings management variables by proposing a novel approach for leveraging the full value of earnings management information as a predictor of corporate failure. We begin by applying a threshold model that separates samples into different regimes, using earnings management as a threshold variable. Earnings management literature identifies two tools used by firms to engage in financial manipulation: accruals and real activities (Chamberlain, Butt, & Sarkar, 2014; Jiang, Zhu, & Huang, 2013). These tools are used in defining earnings management variables. We asymptotically identify three earnings management–level regimes (Regime 1 = low management, Regime 2 = medium management, Regime 3 = high management), in which firms' financial conditions differ with regard to their earnings management. Then, we compare a model based on financial ratios and earnings management information with a traditional corporate failure model, using several prediction methods (logistic regression, support vector machine, neural network, extreme learning machine, and decision tree). Finally, we analyze the impact of earnings management variables on forecasting corporate failure across the regimes derived from the threshold model.

Our study thus makes several contributions. It adds to limited empirical evidence about whether earnings management is an accurate predictor of corporate failure, by presenting a large-sample analysis of earnings-management-based variables as performance drivers of corporate failure models. To the best of our knowledge, our study offers the first assessment of the influence and boundaries of a novel explanatory variable in a corporate failure prediction model. To date, corporate failure studies that have sought to investigate novel explanatory variables have lacked comprehensibility, though such understanding is key to their use in corporate failure models. Accordingly, this study contributes to a novel research area, in which both the ability of explanatory variables to predict corporate failure accurately and the limits of those abilities can lead to superior forecasting of corporate failures.

The remainder of this article is organized as follows: In Section 2, we review prior literature. In Sections 3 and 4, we present our data and research methodology, respectively. Section 5 contains the results, and we conclude in Section 6.

2. Literature review

Although the number of studies devoted to corporate failure forecasting has increased exponentially after the 2008 financial crisis, academics have been developing failure prediction models for more than 60 years. Beaver (1966) published a foundational study of corporate failure prediction in the mid-1960s and used a univariate model to predict corporate failure according to financial ratios. Altman (1968) presented one of the most famous corporate failure measurements, the Z-score, that linearly combines five financial ratios to calculate a score for comparison with a predetermined threshold. The resulting score classifies firms into failed and non-failed.

Today, new knowledge related to the development of corporate failure models has enabled the creation of a wide range of corporate failure prediction methodologies that are more advanced than these pioneering but naive linear models. The development of more sophisticated models that can predict corporate failure is a well-recognized research area. According to Kirkos's (2015) review, in the past decade, most researchers have focused on artificial intelligence and operation research techniques, to develop more accurate, better-performing methods to forecast corporate failure. These techniques have gained popularity and wide application, such as bank failures (Gogas, Papadimitriou, & Agrapetidou, 2018) and prediction markets (Tai, Lin, Chie, & Tung, 2019) because they do not require any statistical assumptions and can model complex nonlinear functions. Researchers use ensemble methods to identify novel techniques that outperform previous techniques; if an artificial intelligence technique provides accurate predictions, combining several different, independent decision rules into one output should provide even better prediction (Kainulainen et al., 2014). According to Verikas, Kalsyte, Bacauskiene, and Gelzinis (2010), bankruptcy prediction research largely relies on two kinds of ensemble methods.

The first method combines the same types of modeling methods, whereas the second combines different types of modeling methods, such as six classifiers using a weighted sum (Sun & Li, 2009). A common element of these models is the assumption that performance depends solely on the quantitative method employed; the models' contributions are determined by the type of quantitative method used.

In addition, researchers have used other methods to develop distinct corporate failure prediction models. Because many factors that predict corporate failure do not appear in financial statements, authors have focused on the predictive power of variables other than financial ratios, such as relational data (Tobback, Bellotti, Moeyersoms, Stankova, & Martens, 2017), audit report information (Muñoz-Izquierdo, Laitinen, Camacho-Miñano, & Pascual-Ezama, 2019), and textual disclosure (Mai, Tian, Lee, & Ma, 2019). These variables complement financial ratio information, resulting in more accurate corporate failure models. Thus, statistical methods that use a variety of explanatory variables can separate failed firms from non-failed firms with the same degree of precision as sophisticated artificial intelligence methods that use financial ratios exclusively.

Since the forecasting performance of corporate failure prediction models can be improved by the use of a variety of explanatory variables, we explore earnings management, a type of information that has seldom been investigated previously. Compared with other explanatory variables that are difficult or costly to gather, relevant only for a given period, or available only for certain types of firms, earnings management information can be calculated from the financial statements of any firm.

To our knowledge, two studies by Serrano-Cinca, Gutiérrez-Nieto, and Bernate-Valbuena (2019) and have explored the use of earnings management as an explanatory variable to improve the prediction of corporate failure. Both studies measure the predictive power of variables according to earnings management information by comparing performance achieved with an earnings management–based model with the performance achieved with a traditional, financial ratio–only model. Both studies support the assertion that earnings management can predict corporate failure. However, the studies measure earnings management differently; whereas Serrano-Cinca et al. (2019) measure it using the financial ratios and total accruals measure proposed by Beneish (1999) to detect financial manipulation, du Jardin, Veganzones, and Séverin (2019) use accrual measures to create earnings management–based variables. In this regard, the Beneish (1999) indicators have raised some concerns because of their weakness in detecting earnings manipulation (MacCarthy, 2017), such that they create significant false positives that misclassify firms as manipulators (Corsi, Di Berardino, & Di Cimbrini, 2015). Nonetheless, these indicators still can be used as early indications of potential financial fraud (Herawati, 2015). Accruals historically have been the primary tool for detecting financial manipulation; if accrual components are properly modeled, abnormal components represent distortions induced by earnings management (Dechow, Ge, & Schrand, 2010). Numerous studies (Enomoto, Kimura, & Yamaguchi, 2015; Jackson, 2018) corroborate the effectiveness of accruals models for predicting failure (Enomoto et al., 2015; Jackson, 2018).

Our study extends these efforts in two ways. First, we examine the predictive power of a new form of data: earnings management information. This study includes *both* tools of earnings management, that is, accruals and real activities. In contrast with du Jardin et al. (2019), who examine only accruals manipulation, we use both tools to establish the overall power of earnings management information to predict corporate failure, reflecting findings from research devoted to financial manipulation, which shows that in addition to using accruals to engage in earning management, firms alter real activities to manage their earnings (Baker, Lopez, Reitenga, & Ruch, 2019; Campa, 2019). Second, our study expands investigations of the role of novel explanatory variables in different corporate failure models. As previously mentioned, recent research includes variables other than financial ratios to build models. However, it focuses only on improving the accuracy associated with novel variables, whereas we use a threshold indicator, based on easy-to-interpret earnings management variables, to evaluate regimes of value in which the prediction capacity of the explanatory variable may vary. Thus, we propose a new methodology for examining the representativeness—the utility and predictive power—of earnings management variables in corporate failure models.

3. Research methodology

In this paper, we extend corporate failure prediction models by proposing a new model based on the conjugation of financial ratios and earnings management-based variables. We also apply a threshold model approach in a novel way, to evaluate the prediction capacities of earnings management variables. Fig. 1 summarizes the proposed prediction framework design.

3.1. Earnings management–based variables

Earnings manipulation can be addressed by detecting accounting abnormalities. Literature that documents such abnormalities has identified two earnings management tools: accruals and real activities manipulation (Campa & Camacho, 2015; Jiang et al., 2013). Because accruals relate directly to earnings, they are subject to being managed to alter earnings, and accordingly, they have been the focus of most empirical research on earnings management—as well as the main component managers use to alter reported earnings. However, managers also manipulate real activities in the course of normal business practices to produce deviations and achieve certain earnings values.

Because the components of earnings manipulation are difficult to identify, researchers approximate models that use components directly related to earnings, that is proxy models (Dechow et al., 2010; Enomoto et al., 2015). Typically, they apply regression to explain the amount of accruals/real activities through the explanatory variables. In that regression, the normal values of the accrual/real activities components are the predicted values from the

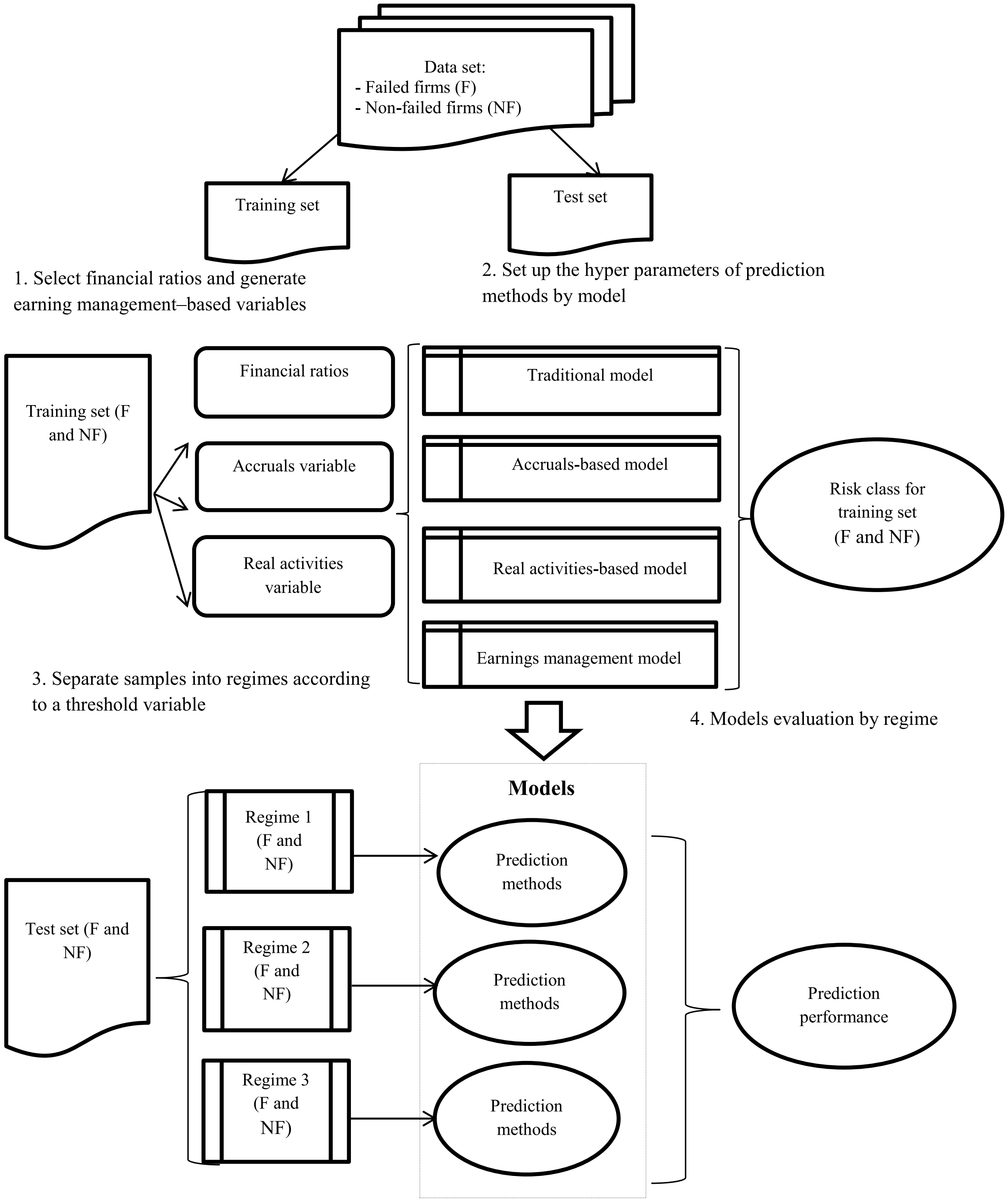

Fig. 1. The methodological design of earnings management-based models based on the threshold model.

regression model, whereas the discretionary accruals/real activities for each firm are represented in the estimation errors. In accounting, regression models have become an accepted method of capturing discretion, because they isolate the managed components of earnings (Dechow et al., 2010). To select proxies, we used two widely accepted models in accounting literature because of their abilities to detect earnings management.

First, we used Kothari, Leone, and Wasley's (2005) model to measure abnormal accruals. This model is an extension of the famous Jones (1991) accruals models, which include return on assets (ROA) to control for the influence of performance on firms' abilities to manipulate accruals. Moreover, du Jardin et al. (2019) support the use of a variable that measure accruals information to predict corporate failure. Kothari et al.'s (2005) accruals model is represented below:

$$\frac{TAcc_{it}}{TA_{it-1}} = \alpha\frac{1}{TA_{it-1}} + \beta\frac{\Delta Sales_{it}}{TA_{it-1}} + \gamma\frac{PPE_{it}}{TA_{it-1}} + \delta\frac{ROA_{it-1}}{TA_{it-1}} + \varepsilon_{it}, \quad (1)$$

where "TAcc" (total accruals) is the change in non-cash working capital before income taxes payable, less total expense depreciation, and the change in non-cash working capital before taxes equals the change in current assets other than cash and short-term investments, less current liabilities other than current maturities of long-term liabilities and income taxes payable. For firm$_i$, TA_{it-1} denotes total assets in year$_{t-1}$; Δ Sales $_{it}$ is sales in year$_t$ less sales in year$_{t-1}$; PPE_{it} refers to gross property, plant, and equipment in year$_t$; ROA_{it-1} denotes net income in year$_{t-1}$;[1] and ε_{it} is discretionary accruals in year$_t$.

Second, we followed Roychowdhury's (2006) methodology to evaluate sales manipulation and overproduction to represent real activities manipulation.[2] Sales manipulation accounts for cash flow to measure whether earnings management has been committed, by accelerating year sales through unusual price reduction, whereas overproduction analyzes production cost manipulation through uncommon applications of scale economies. To capture the total effects of real activities management, we followed Luo, Xiang, and Huang (2017) to aggregate the two measures and compute a comprehensive metric of real-earnings management activities.

The sales manipulation model is represented as follows:

$$\frac{CF_{it}}{TA_{it-1}} = \alpha\frac{1}{TA_{it-1}} + \beta\frac{Sales_{it}}{TA_{it-1}} + \gamma\frac{\Delta Sales_{it}}{TA_{it-1}} + \varepsilon_{it}. \quad (2)$$

The overproduction model is:

$$\frac{Prod_{it}}{TA_{it-1}} = \alpha\frac{1}{TA_{it-1}} + \beta\frac{Sales_{it}}{TA_{it-1}} + \gamma\frac{\Delta Sales_{it}}{TA_{it-1}} + \delta\frac{\Delta Sales_{it-1}}{TA_{it-1}} + \varepsilon_{it}. \quad (3)$$

In these equations, for firm$_i$, CF_{it} defines cash flow in year$_t$; $Prod_{it}$ defines the cost of goods sold plus the change in inventory in year$_t$; TA_{it-1} denotes total assets in year$_{t-1}$; $Sales_{it}$ refers to sales in year$_t$; $Sales_{it-1}$ refers to sales in year$_{t-1}$; $\Delta Sales_{it}$ refers to the change in net sales between year$_t$ and year$_{t-1}$; $\Delta Sales_{it-1}$ refers to the change in net sales between year$_{t-1}$ and year$_{t-2}$; and ε_{it} denotes discretionary earnings in year$_t$.

In our regression, the residual represents a discretionary value of the parameters for each firm. It captures potential earnings management in the residual ε $_{it}$, which we incorporate in our model. Following Hribar and Craig Nichols (2007), we use it to estimate earnings management in absolute values; we capture magnitude using unsigned value error terms. That is, we investigate its impact on corporate failure models according to its magnitude.[3]

3.2. Threshold model

Tong and Lim (1980) introduced threshold models to highlight asymmetric dynamics in cyclical data. Another application of threshold models is sample splitting (Hansen, 2000), which uses the models to estimate the value (threshold) of a variable (predefined using linearity and threshold tests) at which the sample splits into groups (*regimes*), thereby generating asymptotic regimes according to that threshold. This model (a two-stage threshold) is defined as follows:

$$Y_i = \begin{cases} \varnothing_0^{(1)} + \varnothing^{(1)}x_i + \varepsilon_i^{(1)} \ \textit{if} \ Z_i \le s_1 \\ \varnothing_0^{(2)} + \varnothing^{(2)}x_i + \varepsilon_i^{(2)} \ \textit{if} \ s_1 < Z_i < s_2 \\ \varnothing_0^{(3)} + \varnothing^{(3)}x_i + \varepsilon_i^{(3)} \ \textit{if} \ Z_i \ge s_2, \end{cases} \quad (4)$$

where $\varepsilon_i^{(j)}$ $j = 1, 2, 3$; $i = 1, \ldots, n$; white noise independent of variance is represented by $\sigma^2_{(j)}$; n is the number of observations; $x_i = (x_1, \ldots, x_k)'$ are explanatory variables; Z_i represents the threshold variable; and $s_1 ands_2$ represent thresholds 1 and 2, respectively. We used STATA 15 software for the modeling.

We used this application to split the samples according to the earnings management (threshold) variable. Three regimes split the samples (see Appendix A). This finding is in line with Gunny's (2010) documentation of significant differences in the magnitudes of earnings management across sample quantiles. According to the accruals threshold, Regime 1 (low level of manipulation) consists mainly of non-failed firms, and Regime 3 has more failed firms (high level of manipulation). In contrast, according to the real activities threshold, the distributions are opposite and not as pronounced as those of accruals; failed firms slightly outnumber non-failed firms in Regime 1, and the proportion of non-failed firms is slightly higher in Regime 3. We thus note a trade-off between accruals-based and real activities-based earnings management tools, according to firms' financial situations.

4. Data and variables

4.1. Data set description

We collected the data for this study from the financial statements of the Diane French corporate database, which provides financial data for French firms that are required by law to file their annual reports with the French commercial court. The data cover small-and-medium-sized enterprises (SMEs), defined by the following conditions (Campa & Camacho, 2015): (1) 10–250 employees, (2)

[1] We estimate ROA using net income rather than net income plus net-of-tax interest expense (a traditional measure) to address potential problems associated with estimating the tax rate.

[2] Roychowdhury's (2006) methodology evaluates the components of sales manipulation, overproduction, and discretionary expenditures. However, in line with Campa and Camacho (2015), who investigate earnings management in SMEs, we solely focus on sales manipulation and overproduction, because they are the most frequently managed components.

[3] In line with Zang (2012), we winsorize earnings management variables to account for extreme and outlier values that could be caused by noise estimations.

Table 1
Initial financial ratios (financial dimensions).

X1	Account payable/total sales (activity)	X13	Financial expenses/EBITDA (solvency)
X2	Cash/total assets (liquidity)	X14	Financial expenses/net income (solvency)
X3	Cash flow/total debts (profitability)	X15	Long term debts/equity (financial structure)
X4	Cash flow/total sales (activity)	X16	Net income/equity (profitability)
X5	Current assets/current liabilities (liquidity)	X17	Net income/total assets (profitability)
X6	Current assets/total debts (liquidity)	X18	Receivables/total sales (activity)
X7	Current liabilities/total assets (liquidity)	X19	Total debts/EBITDA (financial structure)
X8	EBIT/equity (profitability)	X20	Total debts/total assets (financial structure)
X9	EBITDA/cash flow (profitability)	X21	Total sales/total assets (activity)
X10	EBITDA/total assets (profitability)	X22	Value added/fixed assets (financial structure)
X11	Equity/total assets (financial structure)	X23	Value added/total sales (activity)
X12	Financial debts/equity (solvency)		

turnover of €2–€50 million, and (3) total assets of €2–€43 million. We selected SMEs for this study for two reasons. First, the SMEs represent core firms in industrialized countries, with significant importance for their economies (Ciampi, 2015). Thus, the ability to predict SME failure is crucial for economic vitality. Second, SMEs' financial statements have fewer disclosure obligations—they are less monitored—so they tend to manage earnings better (Ball & Shivakumar, 2005).

Thus, the current study is relevant for two main reasons: First, most studies focus on analyses of listed companies, even though evidence indicates that earnings management is more pervasive in SMEs (Achleitner, Günther, Kaserer, & Siciliano, 2014), which also react differently than listed firms in terms of earnings manipulation behavior than listed firms (Campa & Camacho, 2015). Second, our investigation involves a rarely explored context (i.e., the civil law system) in which earnings management is highly relevant. The French legal system affords less protection to creditors than the common law system (as exists in the United Kingdom or Germany, for example). In such a scenario, French SMEs may have stronger incentives to conduct earnings management practices. Accordingly, this study seeks to enrich knowledge on the influence of earnings management on corporate failure prediction models for French SMEs.

The Diane database provides information about whether firms are failed or healthy; in the case of failed firms, it provides dates of liquidation. A firm is considered to fail once it enters bankruptcy procedure. In our study, we consider a firm healthy if it continued its business activity until at least two years after the collection year; we adopt this conservative perspective to avoid including healthy firms that might suddenly fail and to ensure a reliably non-bankrupt sample. Because we drew our samples from all sectors of activity, we could test their capacities to create good prediction rules. Moreover, we excluded firms with missing values in their financial statements, as well as outliers, to ensure a more stable prediction model. Finally, to facilitate out-of-sample (hold-out) testing, we chronologically divided our data set into a training set and a test set. To control the possible effects of the economic cycle on model performance, we collected multi-year data from 2014 through 2018. During these years, the French economy experienced a period of growth (2014–2015) and a period of a downturn (2017–2018). Therefore, in the first period studied, the financial statements of firms included in the training set were published in 2014, whereas samples in the test set were published in 2015. In the second period studied, firms included in the training and test sets published their financial statements in 2017 and 2018, respectively. With this distinction, we can test the model using data after its construction, thereby ensuring its predictive capacity (Balcaen & Ooghe, 2006). We randomly selected 4000 firms for the training and test sets for the period: 2014–2015 and 4360 firms for the 2017–2018 period. None of the firms in the training set were included in the test set. Both data sets contained the same number of bankrupt and non-bankrupt firms.[4]

4.2. Financial ratios

Ever since Beaver (1966) documented that financial ratios have the power to predict corporate failure, most corporate failure models have used them as explanatory variables. Moreover, financial information reflects several factors that predict failure. Among the vast number of financial ratios that appear in prior research, we focused on factors that have proven to predict SME failure (Cultrera & Brédart, 2016; Gordini, 2014; Kim & Sohn, 2010). We calculated 23 financial ratios specific to SME corporate failure prediction (Table 1). We categorized the ratios according to the five major groups used in financial analysis to examine the financial conditions of a firm: liquidity, solvency, activity, profitability, and financial structure.

Because some ratios possess common numerators or denominators, they may be correlated. Our first step was

[4] The proportion of each type of firms in the data set becomes a significant paradigm in the sample selection process. Because bankruptcy is a rare phenomenon in the real-world population, the proportion of failed firms is low in comparison with that of non-failed firms. To employ a corporate failure model in practice, the sample distribution should be representative of the real-world population (i.e., the data set should be imbalanced). However, this procedure has a significant drawback: The low proportion of failed firms in the entire sample may cause inaccurate decision rules, because the model does not account for the relative distribution of each class (López, Fernández, García, Palade, & Herrera, 2013). To highlight the effect of earnings management variables and avoid the possible effects of an imbalanced data set, we collected a data set in which failed firms represent 50% of the samples. According to Sun, Li, Huang, and He (2014), most research uses balanced data sets; accordingly, we adopt a matched pair technique in which we matched failed firms, according to their industry sector, size and firm age, with non-failed firms. For estimating the conditional probability of failure, we use as much data as possible, while also maintaining a balanced proportion for each sample.

Table 2
Descriptive analysis of selected financial ratios (quartiles).

	Non-failed firms			Failed firms			K-W test
2014–2015 data	25	50	75	25	50	75	
Cash/total assets (X2)	0.06	0.12	0.39	0.01	0.05	0.19	***
Cash flow/total sales (X4)	0.01	0.05	0.12	−0.11	−0.02	0.03	***
Current assets/current liabilities (X5)	1.16	1.77	2.79	0.40	0.76	1.18	***
EBITDA/total assets (X10)	0.02	0.09	0.20	−0.21	−0.04	0.05	***
Total debts/total assets (X20)	0.19	0.40	0.59	0.29	0.52	0.73	***
Value added/total sales (X23)	0.19	0.35	0.53	0.11	0.28	0.44	***
2017–2018 data							
Current assets/current liabilities (X5)	1.02	1.23	2.54	0.32	0.55	1.03	***
EBIT/equity (X7)	0.08	0.11	0.26	−0.19	−0.11	0.03	***
EBITDA/total assets (X11)	0.04	0.10	0.23	−0.14	0.01	0.05	***
Financial expenses/net income (X17)	0.02	0.04	0.05	−0.03	0.07	0.09	***
Receivables/total sales(X18)	0.08	0.09	0.14	0.13	0.16	0.19	***
Total debts/total assets (X20)	0.23	0.56	0.63	0.32	0.65	0.74	***
Total sales/total assets(X21)	1.98	2.43	3.54	1.03	1.98	2.32	***

*** Statistically significant at the 1% level of the difference in medians of the Kruskal–Wallis (K-W) test.

to evaluate the variance inflation factor (VIF). According to Kim, Kang, and Kim (2015), a VIF > 4.0 suggests a risk of multicollinearity. We excluded any variables that exceeded this threshold. However, the number of remaining variables was still too high and could have caused problems of high dimensionalities, such as irrelevancy. Therefore, we needed a variable selection process to choose a subset of the most representative variables with good discriminatory power. Second, we applied a filtering approach that has widely been used in corporate failure prediction. This approach relies on statistical techniques that evaluate the general characteristics of the given data to select the best subset. Finally, we applied stepwise wrapper feature selection (Yu, Miche, Séverin, & Lendasse, 2014) to select the six and seven variables for the first (2014–2015) and the second (2017–2018) data, respectively (Table 2). Notably, the information contained in the financial ratios was completely different from the information contained in the earnings management variables, as indicated by the low value of Pearson's correlation coefficient (see Appendix B).

5. Prediction methods and experimental settings

In this section, we present the five prediction methods used in corporate failure forecasting: logistic regression (LR) (Ciampi, 2015; Jabeur, 2017), support vector machine (SVM) (Barboza, Kimura, & Altman, 2017; Gogas et al., 2018), neural network (NN) (Hosaka, 2019; Petropoulos, Siakoulis, Stavroulakis, & Vlachogiannakis, 2020), extreme learning machine (ELM) (Yu et al., 2014; Zhao, Huang, Wei, Yu, Wang, & Chen, 2017), and decision tree (DT) (Quinlan, 1993). Generally, because no single prediction method uniformly outperforms other methods, the choice of any prediction method depends on its unique pros and cons, as well as the available task. According to Kotsiantis, Zaharakis, and Pintelas (2006), SVMs and NNs tend to perform better when dealing with multiple dimensions, though they require large sample sizes to achieve maximum accuracy. The main advantage of ELMs is their efficiency with regard to computational time/prediction accuracy; it is not necessary to calibrate the unstable parameters because of the random selection of weights and biased values (Huang, Zhu, & Siew, 2004). Although LRs and DTs are easy to implement and are more tolerant of overfitting and irrelevant variables, LRs require a specific distribution of classes for optimal performance; DTs cannot perform well when there is partitioning (Kumar & Ravi, 2007). We selected prediction methods with different characteristics to evaluate whether the methods were particularly sensitive within the regimes calculated. Furthermore, because some prediction methods require a determination of hyperparameters, we describe how we optimized hyperparameters and explain the performance metrics used in the current study.

5.1. Logistic regression (LR)

LR is one of the main prediction methods used in corporate failure models because of its ability to analyze and interpret data and its robustness; it is more suitable than discriminant analysis because it is less demanding about statistical assumptions of prior probabilities and distributions of predictors. The LR method uses nonlinear maximum likelihood to estimate firms' probabilities of failure (Z-scores), then compares them with predetermined thresholds to classify firms in groups. This score is expressed as follows:

$$z = \frac{1}{1 + e^{-\left(\sum_{i=1}^{N} w_i x_i + w_0\right)}}, \tag{5}$$

where x_i denotes explanatory variables, and w_i are weights. They are estimated using maximum likelihood estimation.

5.2. Support vector machine (SVM)

As a popular technique for classification according to a discriminant method to solve complex problems (Vapnik & Vapnik, 1998), SVM relies on the notion of a margin to separate data classes. It maps a training vector into a high-dimensional space, creating maximum distance between the samples at each side of the separating hyperplane, which means maximizing the margin. The instances lying closest to the hyperplane are known as support vectors, and the boundary is represented as a linear combination of those support vectors.

An SVM can be represented according to the following equations

$$Min_{w,e}\frac{1}{2}w^{T}w+C\sum_{i=1}^{N}e_{i} \tag{6}$$

subject to

$$\begin{cases} y_i\left(w\phi\left(x_i\right)+b\right)\geq 1-e_i,\ i=1,\ldots,N \\ e_i\geq 0 \end{cases} \tag{7}$$

where y_i denotes the class value of the training samples, $\phi\left(x_i\right)$ maps training vectors to a high dimensional space; w is the weight vector; b is the bias term; C is the regularization parameter; and e_i is the slack variable.

When the optimal hyperplane separation between classes is built, SVM predicts the data sample using a classifier of the form:

$$f\left(x\right)=sgn\left(\sum_{i=1}^{N}\delta_i y_i K\left(x_i,x\right)+b\right), \tag{8}$$

where *sgn* is the sign function; x is the sample to classify; x_i are the training samples; δ_i is the Lagrange multiplier; and $K\left(x_i,x\right)$ is a kernel function that implicitly maps the input vectors into suitable feature space. One of the most commonly used kernels is the radial basis function (RBF), $K\left(x_i,x\right)=\exp(-\gamma\cdot\|x_i-x\|^2)$, which induces boundaries by placing weighted Gaussians on key training samples (Tong & Koller, 2001). In RBF-SVM, both regularization parameter C and a parameter γ(Gaussian width) must be set up to control for model complexity and training error.

5.3. Neural network (NN)

The NN method, introduced by Messier Jr. and Hansen (1988), has become one of the best-known methods for corporate failure prediction. This method attempts to imitate the human brain, in which interconnected computing units known as neurons are arranged in layers, in an architecture that connects neurons among layers. In our study, the architecture of NN has three layers: an input layer that includes information about the explanatory variables to be processed; an output layer that contains the results of the processing information associated with the classification task; and a hidden layer that contains a certain number of neurons that connect the input and output layers. The strength of the connections among neurons in different layers is represented by numerical weights. This method produces a probability, in this case of failure (Z-score), that can be compared against a threshold to classify the samples, expressed as follows:

$$z=f\left(\sum_{j=0}^{N}w_{lj}f\left(\sum_{i=0}^{d}w_{ji}x_i\right)\right), \tag{9}$$

where f is the activation function; x_i are explanatory variables; w_{ji} is the weights between the input and hidden layers; and w_{lj} is the weights between the hidden and output layers. Our NN uses the Levenberg–Marquardt algorithm as an optimization technique and a hyperbolic tangent as an activation function.

5.4. Extreme learning machine (ELM)

ELM is the latest machine learning technique. Proposed by Huang et al. (2004), ELM is a single hidden layer feedforward neural network (SLFN) in which the weights between the input and hidden layers and bias values are randomly selected. Unlike traditional approaches, ELM avoids issues such as tuning control parameters and/or local minima; it also provides a good analytical solution (Li, Huang, Saratchandran, & Sundararajan, 2005).

Given N distinct samples (x_i, y_i) with $x_i \in R^d$ and $y_i \in R$, the SLFN with k hidden neurons can be modeled as follows:

$$\sum_{i=1}^{k}\beta_i\varnothing\left(w_i x_i+b_i\right),\ 1\leq l\leq N, \tag{10}$$

where $\varnothing$ is the activation function; β_i indicates output weights; and w_i and b_i are the input weights and the biases, respectively. If the SLFN perfectly describes the data, the relationship can be represented compactly as $\boldsymbol{H}\beta=Y$, with

$$\boldsymbol{H}=\begin{pmatrix}\varnothing\left(w_1x_1+b_1\right) & \ldots & \varnothing\left(w_kx_1+b_k\right)\\ \cdots & \cdots & \cdots\\ \varnothing\left(w_1x_n+b_1\right) & \ldots & \varnothing\left(w_kx_n+b_k\right)\end{pmatrix}; \tag{11}$$

$\beta = \left(\beta_1^T,\ldots,\beta_k^T\right)^T$; and $Y = \left(Y_1^T,\ldots,Y_k^T\right)^T$. Thus, the learning procedure initializes the random selection of the w_i and b_i to compute the output weights using the Moore–Penrose pseudo-inverse, $\beta=\boldsymbol{H}^{\tau}Y$.

5.5. Decision tree (DT)

The popularity of the DT arises from its simplicity and capacity to interpret results that provide useful information for classification. Moreover, the DT is computationally efficient in dealing with large numbers of explanatory variables with regard to the number of cases and variable relationships in different parts of the measurement space (Feldman & Gross, 2005). This method generates "trees" that classify samples by recursively splitting input data; the "leaves" represent the corresponding categories, and the "branches" represent the variables that lead to those categories. In this study, we use the algorithm C4.5 developed by Quinlan (1993).

5.6. Hyperparameter optimization

Machine learning algorithms come with a set of hyperparameters that must be tuned to optimize their performance. To this effect, we employed a cross-validation method to test several hyperparameter configurations and to estimate how well they generalized to independent data sets. We conducted this procedure in the following steps:

Step 1. We split the data set into two parts: a training set and a test set. As previously noted (see Section 4.1), we used the hold-out method to divide the data set into training and test sets in chronological order. According to Balcaen and Ooghe (2006), this procedure is common in corporate bankruptcy prediction studies, such that the model can be tested using data after its construction.

Step 2. For each hyperparameter configuration, we applied a 10-fold cross-validation method on the training set to have a training–validation pair for hyperparameter tuning. We performed this cross-validation only for the model selection process, to select the best hyperparameter values. To experiment with various hyperparameter settings, we used a grid search that evaluates each combination of hyperparameters.

Step 3. Taking the hyperparameter settings that achieved the best average result in the cross-validation procedure, we used the complete training set to achieve model fit with these settings. That is, once the hyperparameters have been set, we can leverage the entire training set to learn the actual model.

Step 4. We used the independent test set to evaluate the model obtained in the previous step.

5.7. Performance metrics

In this section, we briefly explain the metrics we used to evaluate the model's performance. First, we used the correct classification rate, which we calculated as follows:

$$\text{Correct classification rate} = \frac{CF + CNF}{TF + TNF}, \tag{12}$$

where CF = correct classification rate of failed firms, CNF = correct classification rate of non-failed firms, TF = total number of failed firms, and TNF = total number of non-failed firms.

Second, we used the failed classification rate and the non-failed classification rate to evaluate the model's predictive capacity for individual failed and non-failed firms, calculated as follows:

$$\text{Failed classification rate} = \frac{CF}{TF}. \tag{13}$$

$$\text{Non} - \text{failed classification rate} = \frac{CNF}{TNF}. \tag{14}$$

To compare the performance of models according to these metrics, we used the Z-score test for two sample proportions (a conventional statistical testing method).

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}\right) + \left(\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}\right)}} \tag{15}$$

where $\hat{p}_1$ and $\hat{p}_2$ are the estimated proportions, n_1 and n_2 are the sample sizes.

Third, we calculated the area under the receiver operating characteristic (ROC) curve (AUC), a widely used measure in the corporate failure prediction domain because it reveals both the false positive rate and the true positive rate according to variance in the decision criteria. It means that the AUC evaluates a model's overall performance without assuming any matrix of misclassification cost. We calculated the statistical differences between the two AUC values using DeLong, DeLong, and Clarke-Pearson's (1988) test.

6. Results and discussion

6.1. Performance comparison between accrual-based and traditional models

We first explored the role of an accruals-based variable on the performance of corporate failure models. Table 3 displays the correct classification rates achieved using solely financial ratios, the rates achieved in combination with the accruals variable for the two periods of time (Panel A and B), and the significant differences (Panel C). These results are represented by the three regimes that we calculate using the threshold model to examine the importance and explanatory power of the accruals variable according to its magnitude. The table shows that the accruals variable enhances the accuracy of the corporate failure model throughout all regimes. Whereas the average correct classification rate achieved with the traditional model was 81.3% and 81.1% in 2015 and 2018, respectively, including the accruals variable, increases the accuracy to 83.2% with data from 2015 and 83.1% with data from 2018 (both significant at the 1% level). This initial result is in line with the findings of du Jardin et al. (2019), who demonstrate the predictive power of this variable. However, the significant improvement in 2015 is mainly the result of better identification of samples that belong to lower- and upper-threshold boundaries. If we analyze the correct classification rates within these two regimes, we note that the accruals-based model achieves increased accuracy, on average, of 3.04 ($p = 0.0104$) and 1.75 ($p = 0.0934$) percentage points in Regimes 1 and 3, respectively. Moreover, the significant improvement obtained in 2018 cannot be attributed particularly to one regime but rather results from a better average correct classification rate among all regimes. These results are reasonable, because the accruals-based model provides better performance, especially with regard to identifying samples that reflect distinct accrual manipulation behaviors. Accordingly, the accruals-based variable is relevant for establishing better discrimination rules.

Tables 4 and 5 show more detail about these initial results, as well as the results of the evaluations of which classes of firms can be better identified. They present classification rates for each class of firm, by correctly classifying failed and non-failed firms. The tables reveal two discrepancies that can explain our previous findings. The first discrepancy relates to the effect of the accruals-based model on the capacity to predict failed and non-failed firms. With regard to all samples of each class (i.e., all regimes), accruals information better identifies failed firms only. In comparison with models based solely on financial ratios, it attains an improvement of 3.6% and 3.4% in 2015 and 2018, respectively (significant at the 1% level). These overall results suggest the relevance of accruals only to failed firms. However, a closer look reveals a significant difference between the regimes in both classes of firms, showing that accruals provide relevant information for both failed and non-failed firms.

Table 3
Correct classification rates: traditional and accruals-based models and significant differences.

2014–2015 data

	Panel A.1: Traditional model (%)						Panel A.2: Accruals-based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	81.6	81.8	82.0	82.4	82.3	82.0	86.4	85.2	84.3	85.4	84.0	85.0
Regime 2	78.2	79.6	79.1	81.3	81.6	80.0	79.2	79.9	80.2	81.4	80.2	80.1
Regime 3	81.7	82.5	79.4	80.6	80.8	81.0	80.7	82.7	81.9	83.7	84.5	82.8
All regimes	81.0	81.7	80.3	81.4	81.5	81.3	82.8	83.2	82.5	84.0	83.5	83.2

2017–2018 data

	Panel B.1: Traditional model (%)						Panel B.2: Accruals-based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	79.6	81.5	81.2	80.4	80.2	80.5	81.7	83.3	83.2	82.1	82.7	82.6
Regime 2	85.4	86.8	83.5	85.3	84.3	85.0	85.6	88.0	85.0	86.6	87.1	86.4
Regime 3	80.7	80.9	81.1	80.2	81.0	80.8	82.2	82.5	83.0	82.7	82.8	82.5
All regimes	80.8	81.9	81.4	81.0	81.1	81.1	82.4	83.7	83.4	83.0	83.3	83.1

Panel C: Significance levels between traditional models and accrual-based model

	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	$+^{***}$	$+^{***}$	$+^{**}$	$+^{**}$	.	$+^{**}$	$+^{*}$	.	.	.	$+^{**}$	$+^{*}$
Regime 2	.	.	.	.	.	.	.	.	.	.	.	.
Regime 3	.	.	$+^{**}$	$+^{***}$	$+^{***}$	$+^{*}$	.	.	.	.	.	.
All regimes	$+^{**}$	$+^{*}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{*}$	$+^{**}$	$+^{**}$	$+^{**}$	$+^{**}$	$+^{**}$

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and an accruals-based variable. Avg. means average. Significance levels are calculated using the Z-test for two sample proportions. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 4
Failed classification rates: traditional and accruals-based models and their significant differences.

2014–2015 data

	Panel A.1: Traditional model (%)						Panel A.2: Accruals-based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	80.1	80.3	82.7	80.8	83.2	81.4	78.0	80.2	78.0	78.8	77.8	78.5
Regime 2	77.6	78.9	80.3	78.2	81.9	79.4	76.9	78.2	78.6	79.6	78.6	78.4
Regime3	78.9	80.1	76.1	77.2	78.6	78.2	86.0	87.1	85.0	87.0	87.4	86.5
All regimes	79.1	80.0	78.6	78.4	80.5	79.3	82.3	83.4	82.0	83.6	83.3	82.9

2017–2018 data

	Panel B.1: Traditional model (%)						Panel B.2: Accruals-based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	78.1	78.9	81.6	79.9	76.8	79.0	79.5	82.0	82.9	80.7	79.7	80.9
Regime 2	81.0	78.1	70.4	74.7	78.1	76.5	84.9	83.5	81.0	82.5	83.9	83.1
Regime3	80.3	79.9	79.4	79.3	78.9	79.5	83.5	83.7	83.1	83.6	83.9	83.5
All regimes	79.2	79.4	79.3	79.0	78.1	79.1	82.2	83.1	82.8	82.5	82.4	82.5

Panel C: Significance level between traditional models and accruals-based model

	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	.	.	$-^{**}$	.	$-^{**}$	$-^{*}$	.	.	.	.	.	.
Regime 2	.	.	.	.	.	.	.	.	$+^{**}$	$+^{*}$	.	.
Regime 3	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{*}$	$+^{**}$	$+^{**}$	$+^{***}$	$+^{***}$	$+^{**}$
All regimes	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{**}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and an accruals-based variable. Avg. means average. Significance levels are calculated using the z-test for two sample proportions. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

The effect of accrual information in the failed firms' regimes is sufficiently strong to suggest that it provides a significant improvement in predicting failed firms overall, whereas the effects on non-failed firms disappear, such that there is no overall improvement.

The explanation of this finding leads to the second discrepancy: Table 4 indicates that the improvement in prediction generated by the introduction of the accruals-based variable is the result of this variable's superior ability to discern failed firms belonging to Regime 3; it has no significant effect on the other two regimes (cf. NN and DT models in Regime 1 in 2015). Table 5 shows that though the accruals variable better identifies non-failed firms in Regime 1, this advantage is annulled by detrimental performance in the upper regime. Therefore, there is a trade-off between the predictive capacity of low and upper regimes among non-failed firms.

Table 5
Non-failed classification rates: traditional and accruals-based models and their significant differences.

	2014–2015 data											
	Panel A.1: Traditional model (%)						Panel A.2: Accruals-based models (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	82.6	82.7	81.7	83.2	81.7	82.4	91.1	88.9	87.9	89.0	87.5	88.9
Regime 2	78.7	80.1	78.2	83.4	81.3	80.3	80.8	81.0	81.3	82.7	81.3	81.4
Regime 3	87.1	87.3	85.6	87.5	85.1	86.5	70.1	73.7	75.5	76.9	78.7	75.0
All regimes	83.0	83.5	82.1	84.5	82.6	83.1	83.3	83.0	83.0	84.3	83.7	83.5
	2017–2018 data											
	Panel B.1: Traditional model (%)						Panel B.2: Accruals-based models (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	80.5	83.0	80.9	80.7	82.3	81.5	83.1	84.1	83.6	82.9	84.5	83.6
Regime 2	87.9	91.3	90.4	90.8	87.7	89.6	86.2	90.6	87.2	89.0	89.0	88.4
Regime 3	82.0	82.9	84.7	82.3	85.5	83.5	79.2	79.9	82.9	80.8	80.3	80.6
All regimes	82.2	84.5	83.6	83.0	84.8	83.6	82.7	84.3	84.1	83.5	84.3	83.8
	Panel C: Significance levels between traditional models and accruals-based models											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	+***	+***	+***	+***	+***	+***	+*	.	+*	.	.	.
Regime 2	.	.	.	.	.	.	.	.	.	.	.	.
Regime 3	−***	−***	−***	−***	−***	−***	.	.	.	.	−**	.
Total	.	.	.	.	.	.	.	.	.	.	.	.

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and an accruals-based variable. Avg. means average. Significance levels are calculated using the z-test for two sample proportions.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

These results are interesting because managers tend to change the images conveyed by their accounts and extensively practice accruals-based manipulation, especially prior to the failure of firms (Habib, Uddin Bhuiyan, & Islam, 2013). The indication that they may intend to use such a practice to mislead, as a last-ditch way to extend their activities, is a signal of their critical situations and a predictor that their firms eventually will fail. Thus, our results indicate that financial institutions should consider such irregularities, avoid lending funds to unsound firms that will fail to meet their obligations. However, the manipulative behavior of sound firms may be detrimental for corporate failure models, because such distortion may lead to misclassifications of sound firms as failed firms. If sound firms decide to take advantage of irregularities, they face the risk that corporate failure models will be unable to predict their fates.

Table 6 presents the results estimated with the AUC measure and the significance levels between AUCs estimated with the accruals-based versus the traditional model. The results with AUCs are consistent with previous results. They confirm that the accruals-based model provides better performance because all AUC differences (for all regimes) are statistically significant.

6.2. Performance comparison between real activities-based and traditional models

Next, we analyze the classification capacities of corporate failure models that introduce a variable based on real activities manipulation. Tables 7 to 10 summarize the results. Specifically, Table 7 displays the overall classification rates; Tables 8 and 9 present the classification rates for failed and non-failed firms, respectively, and Table 10 presents AUC values.

The real activities–based variable leads to a statistically significant improvement in the performance of corporate failure models, but it has less predictive power than the accruals-based variable (Table 3). If we compare the overall correct classification rate for all regimes, the real activities variable provides significant improvements in just three models regardless of the period (only NN in 2015 is significant at 1% level). In contrast, in the model based on the accrual variable, the differences in all correct classification rates are statistically significant (four of ten rates at the 1% threshold). Moreover, the predictive power of the real activity–based variable depends on its capacity to predict which firms extensively embrace this irregularity. Regime 3 exhibits the highest number of statistically significant differences (overall improvement of 3.4% in 2015, significant at the 1% level; improvement of 2.5% in 2018, significant at 5% level), in line with our previous findings that the predictive utility of earnings management variables depends on their magnitudes.

When we examine the overall performance of the real activities–based variable by analyzing its separate effects on failed and non-failed firms, we note that it considerably improves the prediction of non-failed firms. We also note that the improvement occurs mainly in Regime 3, which includes sound firms that have the highest earnings management magnitudes. In contrast, the real activities–based variable has almost no effect with regard to failed firms. Among the 48 differences calculated, only 6 are statistically significant.

The estimated AUCs complement the latter tables, indicating the influence of the real activities-based earnings management measure on model performance. The real activities–based model is significantly better in 11 of 12 cases (6 at a 1% significance level) than those obtained with the traditional model. In particular, improvements

Table 6
Area under the ROC curve (AUC) for traditional and accruals-based models and, their significant differences.

	2014–2015 data											
	Panel A.1: Traditional model						Panel A.2: Accruals-based model					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	0.823	0.826	0.828	0.831	0.820	0.825	0.867	0.860	0.852	0.857	0.837	0.854
Regime 2	0.772	0.786	0.784	0.819	0.821	0.796	0.800	0.803	0.807	0.821	0.802	0.806
Regime 3	0.823	0.836	0.802	0.811	0.809	0.816	0.808	0.832	0.822	0.844	0.851	0.831
All regimes	0.812	0.823	0.810	0.819	0.818	0.817	0.836	0.846	0.845	0.847	0.842	0.843
	2017–2018 data											
	Panel B.1: Traditional model						Panel B.2: Accruals-based model					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	0.802	0.823	0.820	0.813	0.811	0.814	0.823	0.842	0.840	0.832	0.833	0.834
Regime 2	0.862	0.872	0.825	0.849	0.847	0.851	0.871	0.892	0.860	0.878	0.880	0.876
Regime 3	0.811	0.817	0.819	0.809	0.812	0.813	0.840	0.844	0.848	0.842	0.846	0.844
All regimes	0.813	0.823	0.816	0.823	0.821	0.819	0.831	0.852	0.849	0.848	0.850	0.846
	Panel C: Significance levels between traditional model and accruals-based model											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	***	***	*	**	.	**	.	.	.	.	*	.
Regime 2	.	.	.	.	.	.	.	.				.
Regime 3	.	.	.	**	***	.	.	*	*	**	**	**
All regimes	**	***	***	***	***	***	**	***	***	***	***	***

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and an accruals-based variable. Avg. means average. Significance levels are calculated using DeLong et al.'s (1988) test.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 7
Correct classification rates: traditional and real activities–based models and their significant differences.

	2014–2015 data											
	Panel A.1: Traditional model (%)						Panel A.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	81.2	82.0	80.4	80.8	81.2	81.1	82.4	82.2	82.7	81.9	82.6	82.4
Regime 2	81.4	82.1	80.1	81.0	84.5	81.8	82.1	83.1	84.0	82.2	82.9	82.8
Regime 3	80.4	80.2	80.5	84.5	80.0	81.1	84.5	85.2	86.6	82.1	84.4	84.5
All regimes	81.0	81.7	80.3	81.4	81.5	81.2	82.7	82.8	83.5	82.0	82.9	82.8
	2017–2018 data											
	Panel B.1: Traditional model (%)						Panel B.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	81.1	80.6	79.6	79.4	80.8	80.3	80.7	81.4	81.3	80.2	80.8	80.9
Regime 2	74.2	78.5	82.2	78.9	76.1	77.9	77.3	77.3	81.8	80.1	78.0	78.9
Regime 3	83.5	85.0	83.2	83.8	83.3	83.7	86.2	87.1	86.0	86.2	85.5	86.2
All regimes	80.8	81.9	81.4	81.0	81.1	81.2	82.3	82.9	83.2	82.5	82.2	82.6
	Panel C: Significance level between traditional models and real activities—based model											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	.	.	+**	.	+*	.	.	.	.	.	.	.
Regime 2	.	.	+**	.	.	.	.	.	.	.	.	.
Regime 3	+***	+***	+***	—**	+***	+***	+**	+*	+**	+**	+*	+**
All regimes	+**	.	+***	.	+*	+*	+*	.	+**	+*	.	.

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and a real activities–based variable. Avg. means average. Significance levels are calculated using the z-test for two sample proportions.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

in AUCs occur for firms with high levels of real activities–based earning management. All models present significantly better AUCs in Regime 3 (6 of 12 cases at a 1% threshold).

These results show that though both classes of firms can engage in real activities manipulation, sound firms do so more extensively than unsound firms. The inclusion of this novel variable is relevant for financial institutions because it can predict the fates of firms that are sound but that otherwise would be considered failed. Because sound firms carry fewer costs and more opportunities when manipulating real activities (Zang, 2012), such irregular practices may be an intrinsic characteristic of non-failed firms. Moreover, if failed firms manage earnings through real activities, financial institutions cannot derive the advantage of detecting irregularities. Thus, the behavior of failed firms may be camouflaged by the behavior of sound firms.

Table 8
Failed classification rates: traditional and real activities–based models and their significant differences.

2014–2015 data												
	Panel A.1: Traditional model (%)						Panel A.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	78.9	79.8	79.0	78.6	79.8	79.2	80.4	80.3	81.0	81.4	80.4	80.7
Regime 2	77.2	77.6	74.6	74.6	83.8	77.5	76.8	78.7	84.2	76.5	80.9	79.4
Regime 3	82.1	83.7	80.9	81.3	80.9	81.8	80.9	82.5	83.3	70.7	81.7	79.8
All regimes	79.1	80.0	78.6	78.4	80.5	79.3	80.0	80.4	81.8	79.5	80.6	80.5
	2017–2018 data											
	Panel B.1 : Traditional model (%)						Panel B.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	75.3	75.8	75.8	75.7	75.2	75.5	76.4	76.7	77.2	76.9	76.4	76.7
Regime 2	82.7	81.8	83.0	80.9	78.1	81.3	83.0	80.0	84.3	83.0	80.0	82.1
Regime 3	86.0	85.5	84.8	85.0	84.1	85.1	85.7	82.3	85.8	86.3	84.9	85.0
All regimes	79.5	79.4	79.3	79.0	78.1	79.1	79.9	79.8	80.6	80.4	79.2	80.0
	Panel C: Significance levels between traditional models and real activities–based models											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	.	.	$+^{*}$	$+^{**}$	.	.	.	.	.	.	.	.
Regime 2	.	.	$+^{***}$	.	.	.	.	.	.	.	.	.
Regime 3	.	.	.	$-^{***}$	.	.	.	$-^{*}$	.	.	.	.
All regimes	.	.	$+^{***}$	.	.	.	.	.	.	.	.	.

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and a real activities–based variable. Avg. means average. Significance levels are calculated using the z-test for two sample proportions.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 9
Non-failed classification rates: traditional and real activities–based models and their significant differences.

	2014–2015 data											
	Panel A.1: Traditional model (%)						Panel A.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	83.6	84.4	82.0	83.4	82.9	83.3	84.5	84.3	84.6	82.3	85.1	84.1
Regime 2	85.2	86.2	85.2	86.9	85.2	85.7	86.9	87.2	83.8	87.5	84.8	86.0
Regime 3	79.3	77.9	80.3	86.6	79.5	80.7	86.9	86.9	88.7	89.5	86.1	87.6
All regimes	83.0	83.5	82.1	84.5	82.6	83.1	85.4	85.3	85.3	84.5	85.3	85.1
	2017–2018 data											
	Panel B.1: Traditional model (%)						Panel B.2: Real activities–based model (%)					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	88.5	86.8	84.4	84.1	87.9	86.3	86.2	87.4	86.5	84.4	86.4	86.2
Regime 2	67.4	75.6	81.3	77.0	74.2	75.1	72.6	74.9	79.6	77.5	76.1	76.1
Regime 3	81.5	84.6	82.1	82.8	83.1	82.8	86.7	87.9	86.3	86.2	85.9	86.6
All regimes	82.2	84.5	83.6	83.0	84.8	83.6	84.6	86.0	85.9	84.6	85.1	85.2
	Panel C: Significance levels between traditional models and real activities–based model											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	.	.	$+^{**}$	.	$+^{*}$	.	.	.	.	.	.	.
Regime 2	.	.	.	.	.	.	$+^{*}$	.	.	.	.	.
Regime 3	$+^{***}$	$+^{***}$	$+^{***}$	.	$+^{***}$	$+^{***}$	$+^{***}$	$+^{**}$	$+^{**}$	$+^{**}$	.	$+^{**}$
All regimes	$+^{**}$	$+^{*}$	$+^{***}$	.	$+^{***}$	$+^{**}$	$+^{**}$	.	$+^{**}$	.	.	.

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and a real activities–based variable. Avg. means average. Significance levels are calculated using the z-test for two sample proportions.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

6.3. Results of overall earnings management variables in corporate failure models

Our previous findings indicate that the tool a firm uses to conduct earnings management characterizes its condition. The accruals-based variable better predicts failed firms, and the real activities–based variable better predicts healthy firms. We might conclude that the two tools provide complementary information for corporate failure models. However, firms can shift from accruals to real activities manipulation, or vice versa, depending on managers' objectives, effects, and consequences (Zang, 2012). Firms even might conduct *both* types of earnings management simultaneously. Thus, to analyze the overall role of earnings management variables on corporate failure models, it is crucial to investigate both tools together.

Accordingly, we design a model that includes both types of earnings management variables in conjunction

Table 10
Area under the ROC curve (AUC) with traditional and real activities–based models and their significant differences.

	2014–2015 data											
	Panel A.1: Traditional model						Panel A.2: Real activities–based model					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	0.823	0.834	0.819	0.819	0.827	0.824	0.838	0.834	0.840	0.828	0.841	0.836
Regime 2	0.827	0.836	0.809	0.818	0.839	0.826	0.834	0.839	0.845	0.837	0.840	0.839
Regime 3	0.816	0.809	0.819	0.845	0.803	0.818	0.856	0.865	0.872	0.842	0.861	0.859
All regimes	0.812	0.823	0.810	0.819	0.818	0.816	0.843	0.848	0.858	0.840	0.852	0.848
	2017–2018 data											
	Panel B.1: Traditional model						Panel B.2: Real activities–based model					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	0.821	0.813	0.809	0.803	0.816	0.812	0.817	0.821	0.819	0.813	0.819	0.818
Regime 2	0.765	0.801	0.834	0.798	0.776	0.794	0.785	0.791	0.829	0.818	0.794	0.803
Regime 3	0.843	0.865	0.839	0.847	0.842	0.847	0.876	0.882	0.871	0.873	0.870	0.874
All regimes	0.813	0.823	0.816	0.823	0.821	0.819	0.833	0.840	0.851	0.831	0.837	0.838
	Panel C: Significance levels between traditional models and real activities–based model											
	2014–2015 data						2017–2018 data					
	LR	SVM	NN	ELM	DT	Avg.	LR	SVM	NN	ELM	DT	Avg.
Regime 1	.	.	*	.	.	.	.	.	.	.	.	.
Regime 2	.	.	.	.	.	.	.	.	.	.	.	.
Regime 3	***	***	***	**	***	***	***	*	**	**	**	**
All regimes	***	***	***	**	***	***	**	*	***	.	*	**

Notes: The traditional model is based solely on financial ratios. The accruals-based model is based on financial ratios and a real activities–based variable. Avg. means average. Significance levels are calculated using DeLong et al.'s (1988) test.
. Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level

with financial ratios and compare its accuracy with that of traditional corporate failure models. First, we divide the samples according to accruals and real activities regimes to evaluate whether the possible trade-off between them is a crucial factor in distinguishing failed firms from non-failed firms. Second, we determine the results according to regimes (see Appendix D). Table 11 presents the results, including the correct classification rates for all firms in Panel A and the correct classification rates for failed and non-failed firms, respectively, in Panels B and C. Table 12 complements these insights by displaying statistically significant differences between the traditional corporate failure model and the earnings management–based model. We average the results to establish the contribution of the earnings management–based model more precisely.[5]

We note two particularly relevant results. First, accruals-based and real activities–based variables provide significant complementary information to predict corporate failure; the synergy generated by their combination enables the design of an accurate model. For 2015, the overall correct classification rate according to the traditional model is 81.3%, whereas the overall correct classification rate according to the earnings management–based model is 83.9%—an improvement of 2.7% (significant at 1% level). Similarly, for 2018, the traditional model achieves a correct classification rate of 81.1%, whereas the earnings management–based model obtains an improvement of 2.5%, which is significant at a threshold of 1%. However, the improvement in predictions is not consistent within the regimes; it derives from *extreme* regimes. In line with previous results, our model better discerns firms that behave differently with respect to earnings management practices. On the one hand, there is a significant difference between firms that exhibit low levels of accruals manipulation (Regime 1) and high levels of real activities manipulation (Regime 3); the improvement is 7.4 and 4.5 percentage points in 2015 and 2018, respectively ($p = 0.018$ and $p = 0.046$, respectively).

Moreover, there is a significant difference of 3.7 percentage points ($p = 0.022$) in 2015 and 7.2 percentage points in 2018 ($p = 0.003$) in the opposite regime, that is, between firms with high level of accruals manipulation (Regime 3) and low levels of real activities manipulation (Regime 1). However, these improvements are linked directly to the class of the firm. Whereas the former result reflects a better ability to predict non-failed firms (improvements of 9.6% and 5.1% in 2015 and 2018, respectively, significant at the 5% level), the latter result corresponds to a prediction improvement for failed firms (improvements of 8.7% and 7.3% in 2015 and 2018, respectively, significant at the 1% level). Thus, our second finding indicates that the role of earnings management differs according to firm class. Earnings management variables provide relevant information only when there is a trade-off between the two earnings management tools. In this scenario, failed firms are characterized by accruals irregularities, and non-failed firms are characterized by real activities manipulation. Our results confirm the importance of earnings management measures in predicting corporate failure. Designers of new corporate failure prediction models should take their importance into account (see Table 13).

[5] To maintain conciseness, and noting that 960 results (5 methods × 3 samples × 4 accruals regimes × 4 real activities regimes proportions × 2 corporate failure models × 2 data sets) are not readable, the results present the average of each sample.

Table 11
Correct classification rates of models based on financial ratios and models based on both financial ratios and earnings management variables.

Panel A: Correct classification rates for all firms (%)

	2014–2015 data								2017–2018 data							
	Accruals regimes								Accruals regimes							
	Regime 1		Regime 2		Regime 3		All regimes		Regime 1		Regime 2		Regime 3		All regimes	
Real activities regimes	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M
Regime 1	83.5	83.9	79.6	77.3	79.9	83.6	81.8	84.4	83.7	86.2	83.4	83.4	76.9	82.1	81.1	84.7
Regime 2	81.1	84.1	80.5	80.5	82.6	84.2	81.7	83.9	76.6	74.8	76.9	77.6	81.5	80.9	77.9	76.8
Regime 3	76.5	83.9	73.9	76.1	81.4	78.3	78.4	82.4	77.4	81.9	91.2	87.3	84.0	86.7	82.9	85.0
All regimes	82.1	83.9	78.7	77.6	80.8	82.8	81.3	83.9	80.5	82.4	85.3	83.8	80.8	84.6	81.1	83.6

Panel B: Correct classification rates for failed firms (%)

	2014–2015 data								2017–2018 data							
	Accruals regimes								Accruals regimes							
	Regime 1		Regime 2		Regime 3		All regimes		Regime 1		Regime 2		Regime 3		All regimes	
Real activities regimes	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M
Regime 1	80.3	77.0	76.4	75.2	78.7	87.4	79.4	83.5	79.7	81.3	77.3	78.9	71.4	78.7	75.5	79.8
Regime 2	76.4	74.1	77.8	72.2	76.9	78.2	76.8	76.5	80.3	77.5	77.4	77.4	86.6	87.7	81.5	80.3
Regime 3	80.0	69.2	81.8	81.8	82.9	86.5	82.1	81.7	76.4	75.8	72.0	80.0	85.6	89.2	83.1	85.9
All regimes	79.8	76.0	77.1	75.4	79.2	85.9	79.3	82.2	79.1	79.2	76.7	78.6	79.6	84.6	79.2	82.1

Panel C: Correct classification rates for non-failed firms (%)

	2014–2015 data								2017–2018 data							
	Accruals regimes								Accruals regimes							
	Regime 1		Regime 2		Regime 3		All regimes		Regime 1		Regime 2		Regime 3		All regimes	
Real activities regimes	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M
Regime 1	85.7	88.7	83.1	79.5	82.0	75.2	84.6	85.5	87.1	90.4	90.0	88;2	90.9	94.4	88.4	91.1
Regime 2	83.4	88.9	83.3	88.8	92.5	89.3	86.3	90.7	74.2	73.1	76.5	77.9	76.7	74.7	75.2	74.2
Regime 3	60.3	69.9	71.4	74.2	79.1	66.0	76.2	82.8	79.2	84.3	93.3	88.1	80.1	80.5	82.9	84.2
All regimes	83.4	88.5	80.1	79.4	83.2	76.6	83.1	85.7	81.4	84.5	89.6	86.4	83.4	84.5	83.5	84.8

Notes: T.M. = traditional model based solely on financial ratios, E.M. = earnings management model based on financial ratios and both earnings management variables.

Table 12
Significance levels of models based on financial ratios and models based on both financial ratios and earnings management variables.

2014–2015 data	Accruals regimes											
	Regime 1			Regime 2			Regime 3			All regimes		
Real activities regimes	All.	F.	NF	All.	F.	NF	All.	F.	NF	All.	F.	NF
Regime 1	.	.	**	.	.	.	**	***	.	***	***	.
Regime 2	.	.	.	.	.	.	.	.	.	.	.	.
Regime 3	**	.	**	.	.	.	.	.	.	*	.	**
All regimes	**	.	***	.	.	.	**	***	.	***	***	**
2017–2018 data	**Accruals regimes**											
	Regime 1			Regime 2			Regime 3			All regimes		
Real activities regimes	All.	F.	NF	All.	F.	NF	All.	F.	NF	All.	F.	NF
Regime 1	.	.	*	.	.	.	***	***	.	***	**	.
Regime 2	.	.	.	.	.	.	.	.	.	.	.	.
Regime 3	**	.	**	.	.	*	.	*	.			
All regimes	.	.	**	.	.	.	***	***	.	***	***	.

Notes: All = all firms; F = failed firms; NF = non-failed firms; significance levels are calculated using z-test for two sample proportions. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Additionally, we calculated AUC values by type of model and period to assess the significant levels of the differences between the AUCs estimated with the traditional model and earning management–based model. Table 13 shows a statistically significant difference in 14 of 32 cases, with more than 70% of these cases at thresholds of 0.05 or 0.01. Nonetheless, these significant differences are not equally distributed across regimes but are concentrated in those regimes with high levels of earnings management, one type of earnings management, and low levels of other manipulation practices (either accruals or real activities manipulation). Finally, the proposed earnings management–based model leads to the highest overall AUC. This finding is unique because this model provides less costly errors than those estimated with traditional models. Thus, we confirm the relevance of including earnings management variables to predict corporate failure.[6]

[6] Because model estimations are based on a one-year data gap, we have not included macroeconomic conditions in the data sets. Nonetheless, we earnings management–based variables are significant

Table 13
Area under the ROC curve (AUC) by type of model.

Panel A: AUC rates

	2014–2015 data								2017–2018 data							
	Accruals regimes								Accruals regimes							
	Regime 1		Regime 2		Regime 3		All regimes		Regime 1		Regime 2		Regime 3		All regimes	
Real activities regimes	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M	T.M	E.M
Regime 1	0.851	0.859	0.802	0.786	0.806	0.856	0.824	0.856	0.849	0.878	0.829	0.831	0.783	0.847	0.812	0.842
Regime 2	0.803	0.836	0.807	0.809	0.838	0.852	0.826	0.845	0.769	0.753	0.774	0.773	0.824	0.818	0.795	0.786
Regime 3	0.757	0.843	0.746	0.754	0.808	0.775	0.818	0.824	0.764	0.825	0.909	0.879	0.837	0.883	0.837	0.854
All regimes	0.825	0.849	0.796	0.784	0.814	0.838	0.820	0.844	0.813	0.835	0.851	0.837	0.815	0.854	0.817	0.845

Panel B: Significance levels

	2014–2015 data				2017–2018 data			
	Accruals regimes				Accruals regimes			
Real activities regimes	Regime 1	Regime 2	Regime 3	All regimes	Regime 1	Regime 2	Regime 3	All regimes
Regime 1	.	.	**	***	*	.	**	**
Regime 2	.	.	.	.	.	.	.	.
Regime 3	**	.	.	.	**	.	**	.
All regimes	*	.	*	***	*	.	**	***

Notes: T.M. = traditional model based solely on financial ratios, E.M. = earnings management model based on financial ratios and both earnings management variables. Significance levels are calculated using DeLong et al.'s (1988) test. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

6.4. Discussion and implications

Our results affirm and extend studies by du Jardin et al. (2019) and Serrano-Cinca et al. (2019). They show that corporate failure models that include financial manipulation indicators—an accruals indicator, a real activities indicator, or both—along with financial ratios are more likely to predict corporate failure accurately than models that rely solely on financial ratios. However, they also show clear differentiation among earnings management tools. The accruals-based measure better identifies failed firms, whereas the real activities–based measure better identifies non-failed firms. Moreover, the discrimination decisions generated by earnings management–based variables do not provide improvements with regard to all firms. Rather, these improvements occur mainly in the upper threshold, where firms extensively manipulate earnings.

Therefore, earnings management variables identify firms that incur certain magnitudes of earning distortions, on the basis of their different behaviors. If firms have strong motivations for manipulating reported earnings, their choices of accruals or real activities earnings management depend on their financial situations, because each earnings management tool has different benefits and risks. Accruals manipulation is easy to undertake after the fiscal year when managers know whether earnings management is needed (Zang, 2012). Moreover, the cost of performing accruals manipulation is less than the cost of real activities–based manipulation, which implies a deviation from the optimal economic point that negatively affects future performance and liquidity (Bhojraj, Hribar, Picconi, & McInnis, 2009). However, because real activities manipulation does not imply a departure from accounting standards (Campa, 2019), healthy firms tend to engage in real activities manipulation, because they can bear its cost and camouflage it with economic decisions. However, failed firms, with their critical financial situations, are more prone to using accruals manipulation, with their lower cost. Accordingly, information extracted from accruals manipulation characterizes failed firms, and information extracted from real activities identifies non-failed firms. Finally, because accruals and real activities provide complementary information, the most accurate corporate failure models include both variables.

These findings not only support evaluations of the prediction capacity of earnings management variables but also improve comprehension of the predictive power of new explanatory variables used in corporate failure models. Therefore, our study offers three main implications for corporate failure literature. First, in real-world practice, the utility of corporate failure models depends on their predictive abilities. Financial institutions should use earnings management information to identify failed and non-failed firms. Moreover, if firms extensively manipulate their financial accounts, the predictive abilities of traditional accounting-based corporate failure models may be distorted. Because firms rely on the information conveyed in accounting figures, altering how the presented

in both data sets (2014–2015 and 2017–2018), based on two different economic cycles, providing further confirmation of the robustness of our study findings. As a further check, we analyzed longer time-series data between 2014 and 2018 (12,360 failed firms plus 12,360 non-failed firms, for a total of 24,720 firms). With this data set, we carried out a cross-validation loop (10-fold cross validation repeated 50 times) to estimate average evaluation measures. As recommended by Demsar (2006), to compare classifier performance with multiple data, we used a Wilcoxon signed ranks non-parametric test. As Demšar (2006) and Garcia and Herrera (2008) caution, several tests of repeated pairwise comparisons among algorithms could lead to a loss of control of the family-wise error. Thus, we used Holm's (1979) procedure to compute the adjusted *p*-values for the multiple comparison analysis. Due to space limitations, we do not present these results here (the average evaluation measures and significance levels are in Appendix F). The results are in line with our previous findings in terms of the classification performance improvements and significant variables, confirming the influence of earnings management–based variables on the ability to predict the fate of firms.

accounts lead to misrepresentations of their perceived value—which may increase firm misclassifications. The inclusion of earnings management variables (accruals and real activities) may prevent the loss of model performance that results when firms manipulate their accounting figures. If financial institutions leverage this information to assess default probabilities, managers may be dissuaded from managing their earnings. According to our findings, for nearly bankrupt firms, extensive earnings management fails to hide their critical situations; rather, it is a definitive signal of their fate. When sound firms manage earnings (especially via accruals), they may be misclassified as failed firms that are not able to reimburse their debts, thereby increasing their risk of being cut off from funding. Thus, managers' awareness that earnings management variables are being used in corporate failure models should discourage them from engaging in such practices.

Second, though we observe no significant differences in the overall correct classification rate between accruals-based and real activities–based models (see Appendix E), as previously mentioned, the effect on the ability to discern failed and non-failed firms is not identical. Taking into consideration that the asymmetry of misclassification costs between failed and non-failed firms depends on the risk aversion of the model's user (Frydman, Altman, & Kao, 1985), practitioners may wish to prioritize a combination of accruals-based and real activities–based variables. As a result, an earnings management–based model is useful as it tends to be better identify failed and non-failed firms.

In summary, our paper suggests that SMEs provide incentives for earnings management practices and that earnings management behavior can be detected. Therefore, earnings management behavior is not neutral or separate from financial indicators. This finding is particularly relevant because SMEs tend to be more informationally opaque than listed firms. They have fewer legal obligations regarding their data disclosure and therefore are prone to engaging in earnings management strategies. Our study suggests a way to overcome the loss of predictive model performance that results from SMEs altering the figures of their annual accounts. When a measure intended to encompass earning management is included along with financial information, corporate failure models perform better. This, the evidence we provide, regarding the ability of earnings management variables to help predict SMEs' failure likelihoods, is useful for both SMEs and banks. The SMEs depend mostly on bank financing to carry out their activities, so banks' abilities to assess SMEs' fates ultimately should facilitate SMEs' access to finance and help those SMEs with genuine potential compete and grow.

Third, this study offers a methodological implication by establishing the true value of using threshold models to understand the predictive power and boundaries of new variables. Researchers have tended to ignore this value, despite the importance of establishing conditions in which variables can predict corporate failure. Our procedure thus has relevant recommendations for academia, and we hope it serves as inspiration for studies of other methodologies that might enhance understanding of the capacity of new variables to predict failure.

7. Conclusion

Our study specifies the impact of earnings management variables (i.e., accruals and real activity tools) on forecasting corporate failure. Unlike other studies devoted to particular explanatory variables that address only improved accuracy, our proposed model captures, in a novel way, the variables' forecasting powers and limitations. We propose a threshold model to split samples into regimes according to different estimated thresholds of earnings management variables. Accordingly, this study contributes to the body of knowledge about predictor variables by investigating the boundaries of earnings management variables as corporate failure predictors.

The study shows that improved performance of corporate failure models, due to the incorporation of earnings management variables, varies across regimes of earning manipulation levels and between the two tools used to manage earnings. Including earnings management variables in corporate failure models are particularly effective for identifying firms with high levels of manipulation (Regime 3). The accruals-based earnings management variable better forecasts failed firms with high levels of accruals manipulation (Regime 3), but the real activities–based variable better forecasts non-failed firms in Regime 3. Overall, this study highlights the limitation of earnings management information; it suggests the existence of information that is complementary to financial ratios for forecasting corporate failure. However, this complementarity arises only when earnings management variables surpass certain thresholds, in which case firms relate their choices of particular earnings management tools to their financial situations. Finally, this study shows that including both earnings management tools produces the most accurate models, because the tools provide complementary information.

This study has several limitations. First, it analyzes the role of earnings management only at the imminent moment of failure, even though financial manipulation may begin when firms start to experience financial difficulties. Therefore, it could be extended with a panel transition regression, in which the advantages of such models combine with a dual time-individual dimension. Second, because the design of a threshold model requires the selection of explanatory variables, the variable selection may influence the regimes generated by threshold models. Third, because it focuses on SMEs, the earnings management–based model might not apply to large firms, in which financial statements are supervised by outside auditors and the managing of earnings may be more constrained. Fourth, this study does not compare the classification performance of the proposed model across countries that feature differing regulatory systems. Our study examines French firms regulated by civil law, so earnings management is highly relevant because investor protections tend to be weaker. In this scenario, creditors pay more attention to earnings quality, because they are less protected than debtors. Various regulatory systems may affect this propensity to manage earnings.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are very grateful to the anonymous reviewers for their substantial contributions to improving this article.

Appendix A

See Table A.1.

Appendix B

See Table B.1.

Appendix C

See Table C.1.

Appendix D

See Table D.1.

Appendix E

See Tables E.1 and E.2.

Table C.1
Hyperparameter values.

SVM	–	Kernel : Radial Basis Function
	–	Regularization parameter: {10, …, 100}
	–	Gaussian width: {0.1, …, 100}
NN	–	N° of hidden layer nodes: {5, …, 35}
	–	Learning epochs: 50
ELM	–	N° of hidden layer nodes: {5, …, 35}
DT	–	Confidence factor: {0.005, …, 0.05}
	–	Leaf instances: {1, …, 15}

Note: Different hyperparameter values were tested in each model.

Table A.1
Sample distributions within accruals and real activities regimes.

	Accruals-based regimes			Real activities-based regimes			
2014–2015 data	Regime 1	Regime 2	Regime 3	Regime 1	Regime 2	Regime 3	All regimes
Failed firms	573	299	1128	1482	272	246	2000
Non-failed firms	1017	423	560	1322	297	381	2000
All firms	1590	722	1688	2804	569	627	4000
Notes: Accruals as threshold variable: Regime 1 = $s \leq 0.208$; Regime 2 = $0.208 < s \leq 0.327$; Regime 3 = $s > 0.327$. Real activities as threshold variable : Regime 1 = $s \leq 0.073$; Regime 2 = $0.073 < s \leq 0.139$; regime 3 = $s > 0.139$.							
2017–2018 data	Regime 1	Regime 2	Regime 3	Regime 1	Regime 2	Regime 3	All regimes
Failed firms	790	206	1184	1066	325	789	2180
Non-failed firms	1229	407	544	841	425	914	2180
All firms	2019	613	1728	1807	750	1703	4360

Notes: Accruals as threshold variable (2018): Regime 1 = $s \leq 0.321$; Regime 2 = $0.321 < s \leq 0.394$; Regime 3 = $s > 0.394$. Real activities as threshold variable (2018): Regime 1 = $s \leq 0.147$; Regime 2 = $0.147 < s \leq 0.264$; regime 3 = $s > 0.264$.

Table B.1
Pearson correlation coefficients among financial ratios and earnings management variables.

2014–2015 data

	X2	X4	X5	X10	X20	X23	Accruals	Real Act.
X2	1	.34*	.121**	−.043**	.064**	.074**	−.018*	.069**
X4		1	.35*	−.046**	.014	−.067**	−.029*	.023*
X5			1	.420**	.213**	.236**	−.351**	−.056**
X10				1	−.601**	−.037*	.140**	.215**
X20					1	.574**	−.017	.001
X23						1	−.024*	−.066**
Accruals							1	.072**
Real Act.								1

2017–2018 data

	X5	X7	X11	X17	X18	X20	X21	Accruals	Real Act
X5	1	−.078**	.580**	−.165**	.125**	.095**	.038*	.087**	.021*
X7		1	.008	.141**	−.255**	.144**	.322**	.049**	.092**
X11			1	−.131**	.078**	.070**	.106**	.121**	.062**
X17				1	−.099**	.021*	.095**	−.037*	−.041**
X18					1	.087**	−.059**	.061**	.041**
X20						1	.221**	.069**	.025*
X21							1	.155**	.111**
Accruals								1	.167**
Real Act.									1

* Correlation is significant at the 5% level; ** correlation is significant at the 1% level.

Table D.1
Sample distributions by the accruals and real activities regimes.

	2014–2015 data											
	Accruals regimes											
	Regime 1			Regime 2			Regime 3			All regimes		
Real activities regimes	F	NF	ALL	F	NF	ALL	F	NF	ALL	F	NF	ALL
Regime 1	419	678	1097	270	293	563	793	351	1144	1482	1322	2804
Regime 2	89	181	270	18	22	40	165	94	259	272	297	569
Regime 3	65	158	223	11	108	119	170	115	285	246	381	627
All regimes	573	1017	1590	299	423	722	1128	560	1688	2000	2000	4000
	2017–2018 data											
	Accruals regimes											
	Regime 1			Regime 2			Regime 3			All regimes		
Real activities regimes	F	NF	ALL	F	NF	ALL	F	NF	ALL	F	NF	ALL
Regime 1	439	512	951	119	111	230	508	198	706	1066	821	1887
Regime 2	173	260	433	62	68	130	90	99	189	325	427	752
Regime 3	178	458	636	25	227	252	586	247	833	789	932	1721
All regimes	790	1230	2020	206	406	612	1184	544	1728	2180	2180	4360

Table E.1
Significance levels of the difference between accruals-based model and real activities–based.

	2014–2015 data				2017–2018 data			
	All firms	Failed	Non-failed	AUC	All firms	Failed	Non-failed	AUC
LR	.	Λ^{*}	Ψ^{*}	.	.	Λ^{*}	Ψ^{*}	.
SVM	.	Λ^{**}	Ψ^{**}	.	.	Λ^{***}	.	.
NN	.	.	Ψ^{**}	.	.	Λ^{*}	Ψ^{*}	.
ELM	Λ^{**}	Λ^{**}	.	.	.	Λ^{*}	.	Λ^{**}
DT	.	Λ^{**}	.	.	.	Λ^{***}	.	.

Notes: Significance levels are calculated using z-test for two sample proportions. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.
Λ means in favor of the accruals-based model.
Ψ means in favor of the real activities–based model.

Table E.2
Significance levels of the difference between accruals-based model and earnings management–based model.

	2014–2015 data				2017–2018 data			
	All firms	Failed	Non-failed	AUC	All firms	Failed	Non-failed	AUC
LR	Ϙ*	.	Ϙ**	Ϙ**	Ϙ*	.	Ϙ**	Ϙ**
SVM	.	Λ^{*}	Ϙ***	.	.	Λ^{*}	Ϙ***	.
NN	Ϙ*	.	Ϙ***	Ϙ*	.	.	Ϙ**	Ϙ*
ELM	.	.	Ϙ***	.	.	Λ^{*}	Ϙ***	
DT	Ϙ*	.	Ϙ***	Ϙ*	.	.	Ϙ**	Ϙ**

Notes: Significance levels are calculated using z-test for two sample proportions. . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.
Ϙ means in favor of the earnings management–based model.

Table F.1
Correct classification rate by model.

	Overall classification rate				Failed classification rate			
	TM	AM	RM	EM	TM	AM	RM	EM
LR	79.4	80.3	80.1	80.5	76.2	78.5	75.8	77.4
SVM	81.1	81.9	81.8	82.4	79.5	80.7	79.5	80.6
NN	80.6	81.5	81.2	81.8	77.7	79.3	77.8	78.9
ELM	80.5	81.3	81.3	81.5	78.0	79.6	77.8	79.1
DT	80.3	81.3	80.9	81.5	77.8	79.5	77.9	79.0
Average	80.4	81.2	81.0	81.5	77.8	79.5	77.7	79.0

(*continued on next page*)

Appendix F

See Tables F.1 and F.2.

Table F.1 (*continued*).

	Non-failed classification rate				AUCs values			
	TM	AM	RM	EM	TM	AM	RM	EM
LR	82.6	82.1	84.2	83.5	0.802	0.811	0.807	0.823
SVM	82.8	83.1	84.1	84.0	0.818	0.832	0.830	0.834
NN	83.5	83.7	84.7	84.5	0.810	0.824	0.825	0.829
ELM	83.0	83.0	84.8	83.9	0.809	0.821	0.823	0.828
DT	82.9	83.0	83.9	83.9	0.805	0.823	0.819	0.830
Average	82.9	83.0	84.3	84.0	0.809	0.822	0.821	0.829

Notes: TM = traditional model; AM = accruals-based model; RM = real activities–based model; EM = earnings management–based model.

Table F.2
Significance levels of the difference between models.

	Between TM and AM				Between TM and RM				Between TM and EM			
	OCR	FCR	NFCR	AUC	OCR	FCR	NFCR	AUC	OCR	FCR	NFCR	AUC
LR	***	***	.	**	**	.	***	*	***	***	***	***
SVM	**	***	.	***	**	.	***	***	***	***	***	***
NN	***	***	.	***	**	.	***	***	***	***	***	***
ELM	**	***	.	***	**	.	***	***	***	***	***	***
DT	***	***	.	***	**	.	*	***	***	***	***	***
Average	***	***	.	***	**	.	***	***	***	***	***	***

	Between AM and RM				Between AM and EM				Between RM and EM			
	OCR	FCR	NFCR	AUC	OCR	FCR	NFCR	AUC	OCR	FCR	NFCR	AUC
LR	.	***	***	*	.	**	***	***	.	***	*	***
SVM	.	***	***	.	*	.	**	.	*	***	.	*
NN	.	***	***	.	.	.	**	.	*	***	.	*
ELM	.	***	***	.	.	*	**	**	.	***	***	*
DT	*	***	**	*	.	*	**	**	*	***	.	***
Average	.	***	***	.	.	.	**	**	*	***	.	**

Notes: TM = traditional model; AM = accruals-based model; RM = real activities–based model; EM = earnings management–based model. Significant levels based on Wilcoxon non-parametric test; . Not significant; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

References

Achleitner, A. K., Günther, N., Kaserer, C., & Siciliano, G. (2014). Real earnings management and accrual-based earnings management in family firms. *European Accounting Review*, *23*(3), 431–461.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589–609.

Baker, T. A., Lopez, T. J., Reitenga, A. L., & Ruch, G. W. (2019). The influence of CEO and CFO power on accruals and real earnings management. *Review of Quantitative Finance and Accounting*, *52*(1), 325–345.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *The British Accounting Review*, *38*(1), 63–93.

Ball, R., & Shivakumar, L. (2005). Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting and Economics*, *39*(1), 83–128.

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405–417.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 71–111.

Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, *55*(5), 24–36.

Bhojraj, S., Hribar, P., Picconi, M., & McInnis, J. (2009). Making sense of cents: an examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance*, *64*(5), 2361–2388.

Brédart, X., Séverin, E., & Veganzones, D. (2021). Human resources and corporate failure prediction modeling: Evidence from Belgium. *Journal of Forecasting*, http://dx.doi.org/10.1002/for.2770.

Campa, D. (2019). Earnings management strategies during financial difficulties: A comparison between listed and unlisted french companies. *Research in International Business and Finance*, *50*, 457–471.

Campa, D., & Camacho, M. (2015). The impact of SME's pre-bankruptcy financial distress on earnings management tools. *International Review of Financial Analysis*, *42*, 222–234.

Chamberlain, T. W., Butt, U. R., & Sarkar, S. (2014). Accruals and real earnings management around debt covenant violations. *International Advances in Economic Research*, *20*(1), 119–120.

Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. an empirical analysis of Italian firms. *Journal of Business Research*, *68*(5), 1012–1025.

Corsi, C., Di Berardino, D., & Di Cimbrini, T. (2015). Beneish M-score and detection of earnings management in Italian SMEs. *Ratio Mathematica*, *28*(1), 65–83.

Cultrera, L., & Brédart, X. (2016). Bankruptcy prediction: The case of belgian SMEs. *Review of Accounting and Finance*, *15*(1), 101–119.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, *50*(2–3), 344–401.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 837–845.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, *7*, 1–30.

du Jardin, P. (2009). Bankruptcy prediction models: How to choose the most relevant variables? *Bankers, Markets & Investors*, *98*, 39–46.

du Jardin, P., Veganzones, D., & Séverin, E. (2019). Forecasting corporate bankruptcy using accrual-based models. *Computational Economics*, *54*(1), 7–43.

Dutzi, A., & Rausch, B. (2016). Earnings management before bankruptcy: A review of the literature. *Journal of Accounting and Auditing: Research & Practice*, *2016*, 1–21.

Enomoto, M., Kimura, F., & Yamaguchi, T. (2015). Accrual-based and real earnings management: An international comparison for investor protection. *Journal of Contemporary Accounting & Economics*, *11*(3), 183–198.

Feldman, D., & Gross, S. (2005). Mortgage default: Classification trees analysis. *The Journal of Real Estate Finance and Economics*, *30*(4), 369–396.

Forgione, A. F., & Migliardo, C. (2018). Forecasting distress in cooperative banks: The role of asset quality. *International Journal of Forecasting*, *34*(4), 678–695.

Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing recursive partitioning for financial classification: The case of financial distress. *The Journal of Finance*, *40*(1), 269–291.
Garcia, S., & Herrera, F. (2008). An extension on statistical comparisons of classifiers over multiple data sets for all pairwise comparisons. *Journal of Machine Learning Research*, *9*(12), 2677–2694.
García Lara, J. M., Garcia Osma, B., & Mora, A. (2005). The effect of earnings management on the asymmetric timeliness of earnings. *Journal of Business Finance & Accounting*, *32*(3–4), 691–726.
Gogas, P., Papadimitriou, T., & Agrapetidou, A. (2018). Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, *34*(3), 440–455.
Gordini, N. (2014). A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy. *Expert Systems with Applications*, *41*(14), 6433–6445.
Gunny, K. A. (2010). The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research*, *27*(3), 855–888.
Habib, A., Uddin Bhuiyan, B., & Islam, A. (2013). Financial distress, earnings management and market pricing of accruals during the global financial crisis. *Managerial Finance*, *39*(2), 155–180.
Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, *68*(3), 575–603.
Hassanpour, S., & Ardakani, M. N. (2017). The effect of pre-bankruptcy financial distress on earnings management tools. *International Review of Management and Marketing*, *7*(3), 213–219.
Herawati, N. (2015). Application of beneish M-score models and data mining to detect financial fraud. *Procedia-Social and Behavioral Sciences*, *211*, 924–930.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.
Hosaka, T. (2019). Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Systems with Applications*, *117*, 287–299.
Hribar, P., & Craig Nichols, D. (2007). The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research*, *45*(5), 1017–1053.
Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. *Neural Networks*, *2*, 985–990.
Jabeur, S. B. (2017). Bankruptcy prediction using partial least squares logistic regression. *Journal of Retailing and Consumer Services*, *36*, 197–202.
Jackson, A. B. (2018). Discretionary accruals: Earnings management....or not? *Abacus*, *54*(2), 136–153.
Jiang, F., Zhu, B., & Huang, J. (2013). Ceo's financial experience and earnings management. *Journal of Multinational Financial Management*, *23*(3), 134–145.
Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, *29*(2), 193–228.
Kainulainen, L., Miche, Y., Eirola, E., Yu, Q., Frénay, B., Séverin, E., et al. (2014). Ensembles of local linear models for bankruptcy analysis and prediction, case studies in business. *Industry and Government Statistics*, *4*(2), 116–133.
Kim, M. J., Kang, D. K., & Kim, H. B. (2015). Geometric mean-based boosting algorithm with over-sampling to resolve data imbalance problem for bankruptcy prediction. *Expert Systems with Applications*, *42*(3), 1074–1082.
Kim, H. S., & Sohn, S. Y. (2010). Support vector machines for default prediction of SMEs based on technology credit. *European Journal of Operational Research*, *201*(3), 838–846.
Kirkos, E. (2015). Assessing methodologies for intelligent bankruptcy prediction. *Artificial Intelligence Review*, *43*(1), 83–123.
Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, *39*(1), 163–197.
Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning: A review of classification and combining techniques. *Artificial Intelligence Review*, *26*(3), 159–190.
Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques–a review. *European Journal of Operational Research*, *180*(1), 1–28.
Li, M. B., Huang, G. B., Saratchandran, P., & Sundararajan, N. (2005). Fully complex extreme learning machine. *Neurocomputing*, *68*, 306–314.
López, V., Fernández, A., García, S., Palade, V., & Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences*, *250*, 113–141.
Luo, J. H., Xiang, Y., & Huang, Z. (2017). Female directors and real activities manipulation: Evidence from China. *China Journal of Accounting Research*, *10*(2), 141–166.
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*(6), 159–166.
Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. *European Journal of Operational Research*, *274*(2), 743–758.
Mare, D. S. (2015). Contribution of macroeconomic factors to the prediction of small bank failures. *Journal of International Financial Markets, Institutions and Money*, *39*, 25–39.
Messier Jr., W. F., & Hansen, J. V. (1988). Inducing rules for expert system development: An example using default and bankruptcy data. *Management Science*, *34*(12), 1403–1415.
Micha, B. (1984). Analysis of business failures in France. *Journal of Banking & Finance*, *8*(2), 281–291.
Muñoz-Izquierdo, N., Laitinen, E. K., Camacho-Miñano, M. D. M., & Pascual-Ezama, D. (2019). Does audit report information improve financial distress prediction over altman's traditional Z-score model? *Journal of International Financial Management & Accounting*, http://dx.doi.org/10.1111/jifm.12110.
Petropoulos, A., Siakoulis, V., Stavroulakis, E., & Vlachogiannakis, N. E. (2020). Predicting bank insolvencies using machine learning techniques. *International Journal of Forecasting*, *36*(3), 1092–1113.
Quinlan, J. R. (1993). C4. 5: Programming for machine learning. *Morgan Kauffmann*, *38*(48).
Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, *42*(3), 335–370.
Serrano-Cinca, C., Gutiérrez-Nieto, B., & Bernate-Valbuena, M. (2019). The use of accounting anomalies indicators to predict business failure. *European Management Journal*, *37*(3), 353–375.
Sun, J., & Li, H. (2009). Financial distress prediction based on serial combination of multiple classifiers. *Expert Systems with Applications*, *36*(4), 8659–8666.
Sun, J., Li, H., Huang, Q. H., & He, K. Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems*, *57*, 41–56.
Tai, C. C., Lin, H. W., Chie, B. T., & Tung, C. Y. (2019). Predicting the failures of prediction markets: A procedure of decision making using classification models. *International Journal of Forecasting*, *35*(1), 297–312.
Tobback, E., Bellotti, T., Moeyersoms, J., Stankova, M., & Martens, D. (2017). Bankruptcy prediction for SMEs using relational data. *Decision Support Systems*, *102*, 69–81.
Tong, S., & Koller, D. (2001). Support vector machine active learning with applications to text classification. *Journal of Machine Learning Research*, *2*(Nov), 45–66.
Tong, H., & Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of Royal Statistics Society, Series B*, *42*, 245–292.
Vapnik, V., & Vapnik, V. (1998). *Statistical learning theory*. New York: Wiley.
Verikas, A., Kalsyte, Z., Bacauskiene, M., & Gelzinis, A. (2010). Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: A survey. *Soft Computing*, *14*(9), 995–1010.
Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, *128*, 296–302.
Zang, A. Y. (2012). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, *87*(2), 675–703.
Zhao, D., Huang, C., Wei, Y., Yu, F., Wang, M., & Chen, H. (2017). An effective computational model for bankruptcy prediction using kernel extreme learning machine approach. *Computational Economics*, *49*(2), 325–341.