International Review of Economics and Finance 62 (2019) 1–19

Contents lists available at ScienceDirect

International Review of Economics and Finance

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

Predicting corporate bankruptcy: What matters?

Leon Li [a,*], Robert Faff [b]

[a] Waikato Management School, University of Waikato, New Zealand
[b] UQ Business School, University of Queensland, Australia

ARTICLE INFO

JEL classification:
G33
C51

Keywords:
Regime-switching system
Z-score
Distance to default
Bankruptcy

ABSTRACT

Whether accounting: or market-based information should be employed to predict corporate default is a long-standing debate in finance research. Incorporating a regime-switching mechanism, we establish a hybrid bankruptcy prediction model with non-uniform loadings in both accounting- and market-based approaches to reexamine the issue. We find the following. Creditors should increase the loading on market-based information when large and liquid corporations are considered. Conversely, for companies with incremental information involved in accounting reporting proxied by discretionary accruals, banks could emphasize accounting ratio-based variables more than they are already emphasized. Since managerial discretion in accounting numbers could serve as a tool to bring undisclosed information about the firm to the public, the weight on accounting-based information could be increased for firms with high information asymmetry. In addition, the loading on market-based (accounting-based) information should be increased (decreased) during periods of financial crisis, defined by negative gross domestic product growth.

1. Introduction

The approaches to estimate the probability of bankruptcy of a corporate borrower can be grouped into two main categories: Altman's (1968) Z-score model and Merton's (1974) model (for a summary of the research, see Schuermann, 2005). The former seeks to assign the credit quality of corporate borrowers based on their accounting-based information, whereas the latter predicts corporate failure based on the volatility of their equity prices (i.e., market-based information). Whether accounting- or market-based approaches should be employed to predict the likelihood of firm failure is a long-standing debate in accounting and finance research (for a summary, see Hillegeist, Keating, Cram, & Lundstedt, 2004). Accordingly, some researchers suggest combining various failure prediction models instead of using a single measure (e.g., Das, Hanouna, & Sarin, 2009; Kealhofer, 2003; Kealhofer & Kurbat, 2001; Li & Miu, 2010; Löffler, 2007; Miller, 1998; Mitchell & Roy, 2008). It should also be noted that investors and financial institutions, in practice, incorporate different sources of information to arrive at their own credit risk assessments and rarely opt for only one approach.

This study considers the potential of combining the two alternative approaches into a hybrid model and investigates the optimal *loading* between the two types of information. In brief, this study posits that the loading should not be constant across firms and over time. For instance, we point out that creditors are likely to pay more attention to market-based information when they assess the credit risks of large and liquid corporate borrowers, since this information is more reliable and of higher quality. In contrast, banks tend to place more weight on accounting-based information for corporate borrowers that incrementally release information in their accounting reports. In addition, the benefit of managerial discretion with accounting numbers to convey undisclosed information about the firm to the public could be more pronounced for firms with high information asymmetry. Therefore, in such cases, market participants should

* Corresponding author.
E-mail addresses: leonli@waikato.ac.nz (L. Li), r.faff@business.uq.edu.au (R. Faff).

https://doi.org/10.1016/j.iref.2019.02.016
Received 12 July 2016; Received in revised form 24 February 2019; Accepted 25 February 2019
Available online 2 March 2019
1059-0560/

increase the loading on accounting-based information for firms with higher degrees of information asymmetry. Moreover, equity prices can reflect information instantly, whereas the information of financial statements can suffer from delays. We argue that the advantage of market-based information could be more pronounced, particularly regarding financial crises. Consequently, increasing (decreasing) the loading on market-based (accounting-based) information could improve bankruptcy prediction. In sum, a key question addressed in this study is when creditors should pay more attention to market- or accounting-based information. Moreover, if the optimal loadings of each information type are non-uniform, would a dynamic assignment of loadings enhance the overall predictive power of the credit risk model?

In light of the above discussion, this study employs a regime-switching approach to establish a hybrid bankruptcy prediction model with non-monotonic loadings on market- and accounting-based information. In particular, we establish a framework in which two states are defined for capturing two different bankruptcy forecasting alternatives. Moreover, a key feature of our model is its estimation of the probabilities of the specific state for each firm and moment in time by using the data themselves. This study thus adopts this estimated dynamic probability to serve as the loading of each forecasting technique.

We use the five accounting ratio-based variables involved in Altman's Z-score function to establish the accounting-based bankruptcy prediction approach. The average and standard deviation of the equity price in Merton's model are employed to establish the market-based bankruptcy prediction alternative. Finally, we address and test five variables for non-monotonic loading in each approach: firm size, the trading volume of the firm's stock, discretionary accounting accruals, information asymmetry, and a market crisis dummy. More specifically, we measure firm size as the natural logarithm of total assets and the trading volume of the firm's stock is proxied by the average daily trading volume over one quarter. We estimate discretionary accruals using Healy's (1985) model and employ their absolute value as a proxy for the incremental information involved in a firm's financial reporting.[1] We measure idiosyncratic risk using the capital asset pricing model (CAPM) and adopt it as the proxy of information asymmetry. Finally, the U.S. gross domestic product (GDP) growth rate is employed to define a financial crisis dummy. Our sample consists of 31,415 U.S. firm–quarter observations (6394 failed firms and 25,021 non-failed firms) from 1988 to 2011.

We find most of the accounting ratio- and market-based variables to be statistically significant. The results from our bankruptcy prediction model with non-monotonic loadings further show that the loading for the accounting-based approach is negatively related to firm size, the trading volume of the firm's stock, and the market crisis dummy but positively related to absolute discretionary accruals and idiosyncratic risk (and the opposite is true for the market-based approach). Moreover, compared with models that involve just one type of information and the hybrid model with constant loadings, in- and out-of-sample bankruptcy prediction tests support the superiority of the hybrid model with the non-monotonic loadings proposed in this study.

Overall, our results show that the loading of each forecasting approach (i.e., accounting-versus market-based) is not uniform across company borrowers and over time. More specifically, for large and liquid corporation borrowers, banks should increase the loading of market-based information. In contrast, creditors should pay more attention to accounting-based information for companies with incremental information involved in their accounting reporting and a high degree of information asymmetry. Moreover, during periods of market crisis, the loadings on the market- and accounting-based information should be increased and decreased, respectively.

Our study contributes to the literature in several ways. First, we employ a method that recognizes differences in the importance of the accounting- and market-based variables in bankruptcy prediction and establish a hybrid bankruptcy prediction model with endogenous and non-monotonic loadings in each approach, hence producing results that cannot be observed in a traditional binary data analysis model (i.e., logistic model). Second, we employ a regime-switching system to produce dynamic loading that is determined by five variables and link accounting- and market-based bankruptcy prediction alternatives in a continuous and smooth manner. Taking advantage of the less restrictive research design, our study provides evidence that helps to resolve the debate in prior research concerning the importance of Altman's model (i.e., an accounting-based approach) versus Merton's approach (i.e., a market-based method).

The remainder of the paper is organized as follows. Section 2 reviews related studies and develops the research questions. Section 3 discusses the model specifications and demonstrates that the hybrid bankruptcy model with non-monotonic loadings is appropriate for our study. Section 4 describes the data and the measurement of the variables. Section 5 presents the empirical results. Finally, Section 6 concludes the investigation and identifies several directions for future research.

2. Related studies and research questions

2.1. Accounting-versus market-based information

Accounting-based models are typically built by using a set of accounting ratio variables (for a summary of the research, see Tian & Yu, 2017). The most famous such model is Altman's Z-score. Following Altman's (1968) study, Mensah (1984) indicates that past performance involved in a firm's accounting statements might not be informative in predicting the firm's future performance and thus suggests the necessity of revising accounting ratio-based models. Hillegeist et al. (2004) argue that the ability of accounting information to predict bankruptcy is likely to be limited, given that it is formulated to describe the company's financial condition under the going-concern principle (i.e., assuming it will not go bankrupt). Based on criticisms of accounting-based models and the claim that market prices reflect future expected cash flows and should therefore be more useful in predicting bankruptcy, market-based models have been proposed by Black and Scholes (1973) and Merton (1974). Market-based models have been examined by a number of researchers with regard to assessing the probability of default (e.g., Campbell, Hilscher, & Szilagyi, 2008; Hillegeist et al., 2004; Reisz &

[1] Healy's (1985) model is presented in Section 5.2.

Perlich, 2007; Vassalou & Xing, 2004).

Research has also compared market- and accounting-based bankruptcy prediction models. Although Altman's (1968) Z-score model could suffer from a lack of theoretical underpinning, the validity of Merton's model is also limited by a number of stringent assumptions (e.g., Saunders & Allen, 2002). It is therefore not surprising that the empirical evidence on the relative performance of market-versus accounting-based models is mixed (e.g., Agarwal & Taffler, 2008; Blochlinger & Leippold, 2006; Campbell et al., 2006; Hillegeist et al., 2004; Kealhofer, 2003; Oderda, Dacorogna, & Jung, 2003; Reisz and Perlich, 2004; Stein, 2005). Moreover, even if one model is superior to another, this does not imply that the inferior model should be neglected altogether and it might be possible to combine the models to form an even better one (e.g., Kealhofer, 2003; Kealhofer & Kurbat, 2001; Li & Miu, 2010; Löffler, 2007; Miller, 1998; Mitchell & Roy, 2008).

Accordingly, both accounting- and market-based information should be valuable for bankruptcy prediction. This study thus establishes a hybrid bankruptcy prediction model in which both types of information are considered as credit risk indicators and are utilized simultaneously in explaining the probability of default. Moreover, the key feature of our model is that the weights assigned to the two types of credit risk information are not static and could vary across firms and over time. In brief, we establish a hybrid credit risk model with non-monotonic loadings on each type of information using a regime-varying technique. In addition, we judge the performance of the proposed model by comparing it with a hybrid model with static loadings and models that utilize only one type of information.

2.2. Research questions

This study departs from previous related studies by offering a new perspective on the link between accounting- and market-based approaches and explores the effectiveness of a hybrid model that incorporates information from both methods. We examine whether the hybrid model performs better than a single forecasting technique. This issue merits detailed study because both accounting- and market-based variables should be valuable to market participants. Moreover, bank managers should weigh accounting-based variables more heavily at some times and rely more on market-based variables at others.

Our idea is clear. Even if one forecast model is superior to another, it does not follow that one should neglect the others altogether. Although a number of studies have also pointed out that investors and financial institutions rarely opt for only one approach and suggest combining different sources of information to arrive at credit risk assessments, our argument builds on the issue of the determination of the optimal loading for each of the two types of information when they are incorporated in the model. In particular, we contend the loading is not monotonic and we address five research hypotheses, as follows.

Firm size is often used as a proxy for information availability reflected in the market price. Earlier studies amply document that information for large firms is more available and thus more efficiently priced in the market than it is for small firms (Bhattacharya, 2001; Lee & Choi, 2002). Moreover, it is more difficult to short small firms' stocks because these are often held by individual investors. On the other hand, large firms' stocks are held by institutions and it is much easier to borrow such stock for short-selling purposes. In the presence of short sellers, the market is more efficient, with information quickly and correctly reflected in stock prices for large firms (e.g., Boehmer, Jones, & Zhang, 2008; Diether, Lee, & Werner, 2009; Wu, 2007).[2] Accordingly, we expect to observe a positive association between firm size and the loading on market-based information (with the opposite being true for accounting-based information). Based on this reasoning, we establish our first hypothesis as follows.

H1. The loading on market-based (accounting-based) information is positively (negatively) related to firm size.

Increases in liquidity and trading activity are associated with greater market efficiency (e.g., Boehmer & Kelley, 2009; Chordia, Roll, & Subrahmanyam, 2011, 2008; Chordia, Subrahmanyamb, & Tong, 2014; Hendershott and Riordan, 2011; Roll, Schwartz, & Subrahmanyam, 2007). Therefore, a higher trading volume leads to more efficient market trading, where the stock price is more likely to reflect the true value of a firm and market-based variables are thus more reliable. Accordingly, the second hypothesis is as follows.

H2. The loading on market-based (accounting-based) information is positively (negatively) related to the trading volume of the firm's stock.

Two different types of earnings management have been defined by prior studies: efficient and opportunistic earnings management (Scott, 2000). From the perspective of efficient earnings management, managers employ discretionary accruals to improve earnings informativeness in communicating private information, such as the impact of current economic events or private information (Healy & Palepu, 1993). Therefore, earnings management enhances the information content of accounting information (e.g., Dechow & Skinner, 2000; Hann, Lu, & Subramanyam, 2007; Krishnan, 2003; Sankar & Subramanyam, 2001; Tucker & Zarowin, 2006). On the other hand, the perspective of opportunistic earnings management suggests that managers use their discretion and report earnings opportunistically to maximize their own utility (e.g., Balsam, Bartov, & Marquardt, 2002; Burgstahler & Dichev, 1997; Burgstahler & Eames, 2006). Accordingly, discretionary accruals by managers damage financial reporting quality. If earnings management is efficient, then discretionary accruals (earnings management proxy) will have a positive relation with the loading on accounting-based information; if opportunistic, on the other hand, they will have a negative relation. Since the relation can go either way—positive if earnings management is efficient and negative if opportunistic—the hypothesis is non-directional, as follows.

[2] Wu (2007) finds that a stock's price is closer to its fundamental value when it has a larger short volume. Boehmer et al. (2008) report short sellers are well informed and contribute to price efficiency. Diether, Malloy, and Scherbina (2002) suggest that the stocks of small firms are the hardest to short and the least likely to have traded options.

H3. There is a relation between discretionary accruals and the loading on accounting-versus market-based information.

We argue that the importance of accounting-based information could be contingent upon the level of information asymmetry between firm management and corporate outsiders. Specifically, Subramanyam (1996) suggests that the inherent flexibility in Generally Accepted Accounting Principles offers an effective channel for corporate executives to strengthen the value relevance of reported financial statements. Through these discretionary accruals, management can convey private information to market participants about a firm's future performance that is not fully captured in non-discretionary accruals. Moreover, Dechow and Skinner (2000) indicate that managerial discretion in accounting numbers could be one of the critical avenues managers can use to bring undisclosed firm information to light. Similarly, Sankar and Subramanyam (2001), Krishnan (2003), Tucker and Zarowin (2006), and Hann et al. (2007) find evidence supporting the concept of efficient earnings management because such a strategy strengthens communication between firm management and corporate outsiders. Consequently, we hypothesize that accounting-based information becomes more important for firms with high information asymmetry and we establish the fourth hypothesis as follows.

H4. The loading on accounting-based (market-based) information is positively (negatively) related to the level of information asymmetry.

Given the efficient market hypothesis, all the information is quickly and unbiasedly reflected in current stock prices (Doran, Peterson, & Wright, 2010). On the other hand, the information in financial statements could suffer from a delay in the release and thus become less useful for decision making by market participants, particularly during a financial crisis. Based upon this logic, we establish the fifth hypothesis as follows.

H5. The loading on market-based (accounting-based) information is positively (negatively) related to the occurrence of a market crisis.

Based on the above discussions, we aim to investigate whether (and how) the loading on accounting- and market-based information varies with the five variables—firm size, the trading volume of the firm's stock, discretionary accruals, information asymmetry, and crisis occurrence—and establish five corresponding hypotheses. To test these hypotheses, this investigation develops a hybrid bankruptcy prediction model associated with non-monotonic loadings on accounting- and market-based information. In the next section, we discuss the problems of conventional bankruptcy prediction approaches and then demonstrate why the hybrid bankruptcy prediction models with non-monotonic loadings on accounting- and market-based information established by this study could appropriately solve these problems.

3. Model specifications

3.1. Accounting-based approach (model 1)

This work adopts the five accounting ratio-based variables involved in Altman's Z-score function to establish the accounting-based bankruptcy prediction approach. The accounting ratio-based setting is presented as follows:

$$\begin{aligned} y^*_{it+1} &= cont. + \beta_1 WC_{it}/TA_{it} + \beta_2 RE_{it}/TA_{it} + \beta_3 EBIT_{it}/TA_{it} + \beta_4 MVE_{it}/TL_{it} \\ &+ \beta_5 SALES_{it}/TA_{it} + u_{it+1}, u_{it+1} \sim (0, \sigma), \end{aligned} \tag{1}$$

where the explained variable, y^*_{it+1}, $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, represents the credit quality of firms, where the subscript i denotes the ith firm and t+1 denotes the $(t+1)$th quarter, *WC/TA* is working capital divided by total assets, *RE/TA* is retained earnings divided by total assets, *EBIT/TA* is earnings before interest and taxes divided by total assets, *MVE/TL* is the market value of equity divided by total liabilities, and *SALES/TA* is sales divided by total assets.[3]

Notably, y^*_{it+1} is an unobservable latent variable. What we observe is a dummy variable y_{it+1}, defined as $y_{it+1} = 1$ if $y^*_{it+1} > 0$ (i.e., company i defaults in quarter t+1); otherwise, $y_{it+1} = 0$ (i.e., company i does not default in quarter t+1). Subsequently, if the cumulative distribution of u_{it} is logistic, we have what is known as a logit model and the default probability driven by the accounting-based approach becomes

$$\begin{aligned} P^{ACC}_{it+1} &= prob(y_{it+1} = 1 \,|\, WC_{i,t}/TA_{it}, RE_{i,t}/TA_{it}, EBIT_{i,t}/TA_{it}, MVE_{i,t}/TL_{it}, SALES_{i,t}/TA_{it}) \\ &= \frac{1}{1 + e^{-\left(cont. + \beta_1 WC_{i,t}/TA_{it} + \beta_2 RE_{i,t}/TA_{it} + \beta_3 EBIT_{i,t}/TA_{it} + \beta_4 MVE_{i,t}/TL_{it} + \beta_5 SALES_{i,t}/TA_{it}\right)}}. \end{aligned} \tag{2}$$

3.2. Market-based approach (model 2)

The accounting ratio-based variables described above rely mostly on information obtained from companies' financial statements. A contrasting alternative approach is based on work by Merton and has been enhanced by the company KMV. In particular, the distance-to-default (DD) variable derived from Merton's model is employed to capture market-based information. A firm's DD can be computed by knowing both the current level and volatility of its market equity value.

The calculation of the DD is as follows. If we assume that the equity value (E) has a normal probability distribution, the probability

[3] Undeniably, the term *MVE/TL* is associated with the market value of equity. This study even excluded this variable from the accounting-based approach and our conclusions remain robust.

that it will be less than zero (i.e., the value of the assets falls below the value of the debt) is given by

$$P = \int_{-\infty}^{0} p(E, \overline{E}, \sigma_E) dE, \tag{3}$$

where $p(E, \overline{E}, \sigma_E)$ is the normal probability density function with a mean equal to the current equity value ($\overline{E}$) and a standard deviation equal to the standard deviation of the equity value (σ_E). The integral is equal to the integral of the standard normal distribution (φ) from negative infinity to $-\overline{E}/\sigma_E$ and thus the probability of default is obtained from

$$P = \int_{-\infty}^{-\overline{E}/\sigma_E} \varphi(z)dz = \Phi(-\overline{E}/\sigma_E), \tag{4}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal and the value of $\overline{E}/\sigma_E$ is defined as the DD.

Since the DD is a function of both the mean and the standard deviation of the equity value, this study defines the two market-based variables as follows: *EQAVE* is the natural logarithm of the average of the daily equity price over one quarter and *EQSTD* is the natural logarithm of the standard deviation of the daily equity price over one quarter.[4]

Next, we establish the market-based approach:

$$y^*_{it+1} = cont. + \gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it} + u_{it+1}, u_{it+1} \sim (0, \sigma), \tag{5}$$

$$\begin{aligned} P^{MKT}_{it+1} &= prob(y_{it+1} = 1 | EQAVE_{it}, EQSTD_{it}) \\ &= \frac{1}{1 + e^{-(cont.+\gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it})}}. \end{aligned} \tag{6}$$

It should be noted that, at the end of each forecast quarter horizon, the two market-based variables are updated with the most recent market information.

3.3. Hybrid bankruptcy prediction model with static weights (model 3)

Intuitively, the data for both the accounting ratio- and market-based variables could be captured by introducing a method that consists of a specification that includes all the explanatory variables appearing in both forecasting equations. The following model specification is thus established:

$$\begin{aligned} y^*_{it+1} = cont. &+ \beta_1 WC_{i,t}/TA_{it} + \beta_2 RE_{i,t}/TA_{it} + \beta_3 EBIT_{i,t}/TA_{it} + \beta_4 MVE_{i,t}/TL_{i,t} \\ &+ \beta_5 SALES_{i,t}/TA_{it} \\ &+ \gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it} + u_{it+1}, \quad u_{it+1} \sim (0, \sigma), \end{aligned} \tag{7}$$

$$\begin{aligned} P^{Hybrid}_{it+1} &= prob(y_{it+1} = 1 | WC_{i,t}/TA_{it}, RE_{i,t}/TA_{it}, EBIT_{i,t}/TA_{it}, MVE/TL_{it}, SALES_{i,t}/TA_{it}, EQAVE_{it}, EQSTD_{it}) \\ &= \frac{1}{1 + e^{-\left(cont.+\beta_1 WC_{i,t}/TA_{it} + \beta_2 RE_{i,t}/TA_{it} + \beta_3 EBIT_{i,t}/TA_{it} + \beta_4 MVE_{i,t}/TL_{it} + \beta_5 SALES_{i,t}/TA_{it} + \gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it}\right)}}. \end{aligned} \tag{8}$$

This regression analysis includes the five accounting ratio-based variables and the two market-based ones as the explanatory variables for firm credit quality. Notably, the impacts of the independent variables on the dependent variable (i.e., β_i, $i = 1, 2, \ldots, 5$ and γ_i, $i = 1, 2$) are fixed across firms and time. Accordingly, there exists a weight w to reformulate Equation (8) as follows:

$$\begin{aligned} P^{Hybrid}_{it+1} = w \times & \frac{1}{1 + e^{-\left(cont_1 + \beta^*_1 WC_{i,t}/TA_{it} + \beta^*_2 RE_{i,t}/TA_{it} + \beta^*_3 EBIT_{i,t}/TA_{it} + \beta^*_4 MVE_{i,t}/TL_{it} + \beta^*_5 SALES_{i,t}/TA_{it}\right)}} \\ & + (1 - w) \times \frac{1}{1 + e^{-(cont_2 + \gamma^*_1 EQAVE_{it} + \gamma^*_2 EQSTD_{it})}} \\ & = w \times P^{ACC}_{it+1} + (1 - w) P^{MKT}_{it+1}. \end{aligned} \tag{9}$$

Equation (2) for the accounting-based approach and Equation (6) for the market-based approach are two special cases of Equation (9) that are restricted by $w = 1$ and $w = 0$, respectively. We thus denote the model specifications in Equations (2) and (6) as the model for the accounting- and market-based approaches, respectively.

[4] To be consistent with Equation (4), this study adopts the standard deviation (i.e., volatility) of stock prices instead of returns. For a more general construction of the DD, see Bharath and Shumway (2008).

3.4. Hybrid bankruptcy prediction model with non-monotonic weights (model 4)

Model 3 is limited by static weights. In our final model, we adopt a state-varying system to establish the following model specification:

$$y^*_{it+1} = \begin{cases} cont_1 + \beta_1 WC_{i,t}/TA_{it} + \beta_2 RE_{i,t}/TA_{it} + \beta_3 EBIT_{i,t}/TA_{it} + \beta_4 MVE_{i,t}/TL_{it} \\ \qquad +\beta_5 SALES_{i,t}/TA_{it} + u_{it+1}, \text{if} \quad s_{it+1} = 1, \\ cont_2 + \gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it} + u_{it+1}, \text{if} \quad s_{it+1} = 2. \end{cases} \tag{10}$$

Notably, s_{it} is a state variable and a two-state system is defined in Equation (10). Regime I (namely, $s_{it+1} = 1$) is set when firm credit quality (i.e., y^*_{it+1}) depends on accounting ratio-based variables, while regime II (namely, $s_{it+1} = 2$) is set when we use market-based variables to measure firm credit quality. Next, to control the dynamic process of the state variable s_{it}, we make the probability conditional on several information variables:

$$w_{it+1} = prob(s_{it+1} = 1|\Pi_{it}) = \frac{\exp(\theta_0 + \theta_1 \cdot \pi_{1,it} + \theta_2 \cdot \pi_{2,it} + \theta_3 \cdot \pi_{3,it} + \cdots)}{1 + \exp(\theta_0 + \theta_1 \cdot \pi_{1,it} + \theta_2 \cdot \pi_{2,it} + \theta_3 \cdot \pi_{3,it} + \cdots)}, \tag{11}$$

$$1 - w_{it+1} = prob(s_{it+1} = 2|\Pi_{it}) = 1 - prob(s_{it+1} = 1|\Pi_{it}), \tag{12}$$

where Π_{it} is a $K \times 1$ vector of variables for controlling the state probabilities.

Although the state variable (s_{it+1}) is unobservable, the data can be used to estimate the probability of a specific regime for a certain firm at any given time. This work thus applies the estimated probability to serve as the loading for each technique. The estimated default probability of firms can thus be presented as follows:

$$\begin{aligned} P^{Hybrid}_{it+1} = w_{it+1} \times & \frac{1}{1 + e^{-\left(cont_1 + \beta_1 WC_{i,t}/TA_{it} + \beta_2 RE_{i,t}/TA_{it} + \beta_3 EBIT_{i,t}/TA_{it} + \beta_4 MVE_{i,t}/TL_{it} + \beta_t SALES_{i,t}/TA_{it}\right)}} \\ & + (1 - w_{it+1}) \times \frac{1}{1 + e^{-(cont_2 + \gamma_1 EQAVE_{it} + \gamma_2 EQSTD_{it})}} \\ & = w_{it+1} \times P^{ACC}_{it+1} + (1 - w_{it+1}) \times P^{MKT}_{it+1}. \end{aligned} \tag{13}$$

A comparison of Equations (9) and (13) clearly reveals the difference between the two hybrid models, distinguished by static versus varying weights. Section 5.6 examines this difference more closely. Importantly, the probabilities of the specific state, namely, $prob(s_{it+1} = 1 \text{ or } 2 \mid \Pi_{it})$, are estimated by the data themselves and will change across firms and over time. In the following discussion, the model specifications for Equations (10)–(13) are applied to a hybrid model with non-monotonic weights and labeled as Model 4.[5]

4. Sample and variable measures

4.1. Sample

We identify failed firms over the period 1988–2011 using the two criteria available in the Compustat database: the status alert (STALTQ) and the reason for deletion (DLRSN). By definition, firms with STALTQ = TL are firms involved in bankruptcy court proceedings and DLRSN = 02 and DLRSN = 03 define firms deleted from the database due to bankruptcy and liquidation, respectively. We use the time of the status alert and deletion as the time of each default event. We collect the data on a quarterly basis, employing data in the estimations until the last available quarter before the time of the default event.[6] For each failed firm, we select the three largest firms (defined by the value of total assets) in the same industry as in the bankruptcy year—defined by four-digit Standard Industrial Classification (SIC) codes—and these firms constitute the non-failed firm sample.[7] Following convention, we exclude financial firms (i.e., SIC codes 6000–6999),[8] non-U.S. firms (including American depositary receipts and overseas branches), firms classified as part of the non-operating establishment industry (i.e., SIC code 9995), firms associated with missing data, and firms from industries without enough

[5] We estimate the model parameters using the maximum likelihood estimation method. See the Appendix for details.

[6] Since the failed firms have a problem with missing data, the forecast horizon is not perfectly consistent in the sample of defaulting firms. For example, the firm ABS Industries, Inc was deleted from the database for default on December 31, 1996 (1996:Q4). The last available data for the firm are dated July 31, 1995 (1995:Q3). The default time for the firm Aero Systems Inc is May 31, 1994 (1994:Q2) and the last available data for the firm are dated November 30, 1992 (1992:Q2).

[7] Using large firms in the same industry, which are likely of better credit quality, as our non-failed firms ensures that our samples of failed and non-failed firms are very different from each other in terms of credit quality. Large firms are deemed to be in good financial health and far from experiencing any distress. The discriminatory power of the models examined in this study can therefore be enhanced, enabling us to achieve a more refined comparison of the performances of alternative models.

[8] The literature on bankruptcy predictability has tended to exclude financial firms from samples because their financial statements are very different from those of companies in most other industries, given the high concentration of financial assets and liabilities.

non-failed firms to match the failed ones. Accordingly, the final sample of failed firms consists of 6394 firm–quarter observations from 421 individual firms.[9] For the non-failed firms, we have 25,021 firm–quarter observations obtained from 441 individual firms.[10] We obtain financial statement information from the Compustat database. Information on equity prices is obtained from the Center for Research in Security Prices database.[11]

4.2. Measures of variables

Our empirical analyses are based on a regression of a bankruptcy dummy (equal to one if the firm is bankrupt and zero otherwise) on five accounting ratio– and two market-based variables (all defined in Section 3). The key feature of this study is to address five variables to control for the non-uniform loadings on accounting- and market-based information and develop the corresponding hypotheses (discussed in Section 2). We measure firm size (*SIZE*) as the natural logarithm of total assets. Trading volume (*VOLUME*) is proxied by the average daily trading volume (in millions of dollars) over one quarter. Next, we define a dummy variable for market crises (*CRISIS*) using a negative value of the U.S. GDP growth rate.

Graham et al. (2005) indicate that earnings management is pervasive but not always observable by market participants. We therefore use discretionary accruals as a proxy for earnings management (e.g., Cheng and Warfield 2005; Bergstresser and Philippon 2006; Larcker et al., 2007). Because reported earnings can be manipulated either upward or downward, we further adopt the absolute value of discretionary accruals ($|DA|$) to gauge the degree of earnings management.

In this study, we estimate the discretionary accruals (*DA*) using Healy's (1985) model. Specifically, we first estimate the following equation[12]

$$TACC_\tau = (NI_\tau - OCF_\tau)/TA_{\tau-1}, \tag{14}$$

where.

$TACC_\tau$	=	total accruals deflated by prior period assets in year τ,
NI_τ	=	net income after taxes in year τ,
OCF_τ	=	operating cash flow in year τ,
$TA_{\tau-1}$	=	total assets at τ - 1 year-end.

Next, we use the average of the previous *TACC* value to compute non-discretionary accruals (*NDA*), as follows:

$$NDA_\tau = \sum_{t=1}^{T} \frac{TACC_t}{T}, \tag{15}$$

where.

NDA_τ	=	estimated non-discretionary accruals in year τ,
t	=	subscript indicating a year included in the estimation period,
τ	=	subscript indicating a year in the event period.

To test H4 regarding the issue of information asymmetry, we adopt the empirical application of the CAPM as follows:

$$R_{i,t} = \alpha_i + \beta_i \times R_{m,t} + e_{i,t}, \tag{16}$$

where $R_{i,t}$ and $R_{m,t}$ denote the daily excess returns of the *i*-th stock and broad market portfolio at time *t*, respectively. The idiosyncratic risk (*IRISK*) of a stock is computed as the standard deviation of the regression residuals from Equation (16). We employ *IRISK* as a measure of a firm's information asymmetry (e.g., Barry & Brown, 1985; Dierkens, 1991; Moeller et al., 2007).

[9] Because of the limited number of bankruptcy observations, researchers have suggested adopting the credit spread and credit default swap spread to examine the issue of credit risk (e.g., Castagnetti & Rossi, 2013; Das et al., 2009; Koopman & Lucas, 2005). It should be noted that the credit spread and credit default swap spread can only capture losses due to downgrades and these are relatively insignificant compared to losses in credit risk, since the latter depend mostly on extreme losses due to bankruptcy and not on downgrades.

[10] It should be noted that some industries have fewer than three non-defaulted firms in the default year. Therefore, we might select the same non-failed firm to match the failed year.

[11] For information used to measure the credit risk of retail customers, see Watkins, Vasnev, and Gerlach (2014).

[12] Some researchers suggest estimating discretionary accruals using Jones's (1991) model and the modified Jones model after controlling for prior performance, as Kothari, Leone, and Wasley (2005). Since the missing data issue is severe for some failed firms in this study, we adopt Healy's (1985) model for *DA* estimation. It should be noted that the various measures of discretionary accruals are highly related. Moreover, Balsam (1998, p. 231, Table 1) indicates that the Jones and Healy models exhibit a high correlation of 0.84.

Table 1
Descriptive statistics of variables.

	Failed firms			Non-failed firms			Test of equal means	
	N	Mean	S.D.	N	Mean	S.D.	*t-value*	*p-value*
Accounting-based variables								
WC/TA	6394	0.213	0.226	25,021	0.187	0.166	10.163	0.000
RE/TA	6394	−0.243	0.881	25,021	0.225	0.374	−64.289	0.000
EBIT/TA	6394	−0.006	0.049	25,021	0.025	0.023	−73.971	0.000
MVE/TL	6394	2.919	5.308	25,021	2.765	3.625	2.725	0.000
SALES/TA	6394	0.359	0.231	25,021	0.333	0.204	9.002	0.000
Market-based variables								
EQAVE	6394	1.855	1.020	25,021	3.190	0.876	−105.09	0.000
EQSTD	6394	2.245	3.093	25,021	1.054	1.564	29.821	0.000
Determinant variables of optimal loadings								
SIZE	6394	4.936	1.592	25,021	7.749	1.738	−117.43	0.000
VOL	6394	10.653	1.690	25,021	12.855	1.942	−82.987	0.000
$\|DA\|$	6394	0.072	0.096	25,021	0.045	.0489	30.992	0.000
IRISK	6394	0.048	0.026	25,021	0.024	0.015	93.869	0.000
GDP	6394	0.013	0.005	25,021	0.012	0.007	13.875	0.000

Variable definitions.
WC/TA = Working capital/Total assets.
RE/TA = Retained earnings/Total assets.
EBIT/TA = Earnings before interest and taxes/Total assets.
MVE/TL = Market value of equity/Total liabilities.
SALES/TA = Sales/Total assets.
EQAVE = The natural logarithm of the mean of daily equity price over one quarter.
EQSTD = The natural logarithm of the standard deviation of daily equity price over one quarter.
SIZE=The Natural logarithm of the book value of total asset.
VOL=The Natural logarithm of the mean of daily stock trading volume over one quarter.
$|DA|$ = Absolute discretionary accruals (Total accruals less nondiscretionary accruals).
IRISK=The standard deviation of the regression residuals from the CAPM.
GDP = U.S. GDP growth rate.
The sample consists of 31,415 firm-quarter observations (6394 for failed firms and 25,021 for non-failed firms) for the period from 1988 to 2011.

5. Empirical results

5.1. Descriptive statistics

Table 1 summarizes the descriptive statistics of the five accounting ratio-based variables, two market-based variables, and five determinant variables for the non-uniform loadings on accounting- and market-based information (i.e., firm size, the trading volume of the firm's stock, absolute discretionary accruals, idiosyncratic risk, and U.S. GDP growth rate). Moreover, a comparative analysis between failed and non-failed firms is listed. As shown in Table 1, the inequality between failed and non-failed firms is significant for all the variables. Table 2 lists the correlation coefficients of all the variables. A high correlation coefficient is only observed between *SIZE* and *VOL* (correlation = 0.805), as well as between *EQAVE* and *EQSTD* (correlation = 0.803).[13]

5.2. Estimation results of alternative bankruptcy prediction specifications

Table 3 lists the estimation results of the two bankruptcy prediction models involving just one type of information. Panel A shows the results of the accounting-based approach (Model 1), while those of the market-based approach (Model 2) are presented in Panel B. With the 5% significance level as a criterion, the estimates of the effects of the accounting ratio-based variables on a firm's bankruptcy probability are significant. Next, as shown in Panel B, *EQAVE* and *EQSTD* (the two market-based factors) are significantly negative and positive, respectively. This result is consistent with the notion that failed firms are associated with a lower equity value and higher equity volatility.

Table 4 presents the estimation results of the hybrid bankruptcy prediction model with static weights (Model 3). First, the accounting ratio-based variables are significant, except for *WC/TA*. Next, *EQAVE* and *EQSTD* are significantly negative and positive, respectively.

[13] High intercorrelations among explanatory variables are neither necessary nor sufficient for a multicollinearity problem. The best indicators of the problem are revealed by testing the stability of the estimated coefficients when some observations are deleted (Maddala & Lahiri, 2009). In Section 5.5, we conduct an out-of-sample test by randomly dropping 10% of the samples and rerunning Model 4. The estimation results are presented in Table 7. We find qualitatively the same results as those reported in Table 5. As shown, the multicollinearity issue should not seriously affect our inferences.

Table 2
Correlation matrix of variables.

Variables	*WC/TA*	*RE/TA*	*EBIT/TA*	*MVE/TL*	*SALES/TA*	*EQAVE*	*EQSTD*	*SIZE*	*VOL*	*|DA|*	*IRISK*	*GDP*
WC/TA	1.000											
RE/TA	0.075	1.000										
EBIT/TA	0.016	0.485	1.000									
MVE/TL	0.339	0.029	0.146	1.000								
SALES/TA	0.159	0.097	0.181	−0.022	1.000							
EQAVE	−0.065	0.484	0.494	0.188	−0.025	1.000						
EQSTD	−0.028	0.335	0.345	0.197	−0.015	0.803	1.000					
SIZE	−0.313	0.330	0.332	−0.105	−0.145	0.653	0.493	1.000				
VOL	−0.167	0.133	0.223	0.112	−0.104	0.447	0.433	0.805	1.000			
|DA|	0.095	−0.202	−0.179	0.093	0.064	−0.168	−0.099	−0.217	−0.097	1.000		
IRISK	0.047	−0.383	−0.402	−0.023	0.035	−0.646	−0.250	−0.519	−0.305	0.190	1.000	
GDP	0.036	−0.022	−0.006	0.042	0.026	−0.020	−0.060	−0.123	−0.169	0.025	−0.002	1.000

The sample consists of 31,415 firm-quarter observations (6394 for failed firms and 25,021 for non-failed firms) for the period from 1988 to 2011. The definition of the variables is consistent with Table 1.

Table 3
Estimation results of the alternative bankruptcy prediction models.

Panel A: Accounting-based approach (Model 1)				
	Coefficient	S.D.	t-value	p-value
Constant	−1.6487	0.0337	−48.9100	0.0000
WC/TA	0.8153	0.0942	8.6600	0.0000
RE/TA	−1.2363	0.0412	−30.0100	0.0000
EBIT/TA	−27.2388	0.6610	−41.2100	0.0000
MVE/TL	0.0289	0.0046	6.2500	0.0000
SALES/TA	1.5595	0.0717	21.7600	0.0000
Log-likelihood	−12812.9			
Panel B: Market-based approach (Model 2)				
	Coefficient	S.D.	t-value	p-value
Constant	2.8573	0.0727	39.3300	0.0000
EQAVE	−1.6337	0.0287	−56.9700	0.0000
EQSTD	0.4238	0.0288	14.7000	0.0000
Log-likelihood	−11723.3			

Table 4
Estimation results of the hybrid bankruptcy prediction model with static weights (Model 3).

	Coefficient	S.D.	t-value	p-value
Constant	1.6596	0.0861	19.2600	0.0000
Accounting-based variables				
WC/TA	0.0716	0.0989	0.7200	0.4690
RE/TA	−0.2409	0.0370	−6.5000	0.0000
EBIT/TA	−15.1902	0.6516	−23.3100	0.0000
MVE/TL	0.0852	0.0044	19.3900	0.0000
SALES/TA	1.3688	0.0804	17.0300	0.0000
Market-based variables				
EQAVE	−0.9798	−1.3610	0.0324	−42.0500
EQSTD	2.2639	0.2832	0.0301	9.4200
Log-likelihood	−11062.8			

This table presents the estimation results of the hybrid bankruptcy prediction model with static weights (see Equations (7) and (8) for the details of the model specifications). See Table 1 for variable definitions and sample descriptions.

Finally, the last row of Table 4 lists the value of the log-likelihood function. Model 3 is associated with a higher value of the log-likelihood function in comparison with Models 1 and 2. The result indicates that the hybrid model in which both the accounting ratio- and market-based variables are included better explains the in-sample variation of the dependent variable (i.e., bankruptcy or liquidation events).

Lastly, Table 5 presents the estimation results of the hybrid bankruptcy prediction model with non-monotonic weights established by this study (Model 4). First, all five accounting ratio-based variables are significant at the 1% level. Second, *EQAVE* and *EQSTD* are significantly negative and positive at the 1% 1evel, respectively. Next, we examine the effect of the five determinant variables on the non-uniform loadings addressed by this study. First, the coefficients of *SIZE* and *VOL* are negative and significant at the 1% level. Second, the coefficients of $|DA|$ and *IRISK* are significantly positive at the 1% level. Third, the coefficient on *CRSIS* is negative and significant (coeff. $= -0.6711$ and p-value $= 0.0000$). Importantly, Model 4 (the hybrid bankruptcy prediction model with non-monotonic weights) has the highest value of the log-likelihood function in comparison with Models 1 to 3.

5.3. Implications and discussion

The key issue in this study is the examination of the non-uniform loadings for each forecasting technique. We assume a state-varying system in which an unobservable state variable with possible outcomes of one and two is designed and the two states are linked via a probability conditional on the five determinant variables, namely, firm size, the trading volume of the firm's stock, discretionary accounting choices, information asymmetry, and crisis occurrence. As shown in Table 5, both *SIZE* and *VOL* are negative and significant. This result implies an increasing (decreasing) weight of market-based (accounting-based) information for large and liquid firms. Since the market-based information for large and liquid firms is more qualified and reliable than that for small and illiquid ones, market-based information becomes more important in explaining default risk in comparison with accounting-based information. Our first two hypotheses (H1 and H2) are supported by the negative estimates of *SIZE* and *VOL*.

The positive coefficient of $|DA|$ implies a relatively high (low) loading of accounting-based (market-based) information for firms associated with high discretionary accounting choices. This result is consistent with the notion that the discretionary accounting choices of firms include incremental information content (i.e., the perspective of efficient earnings management) and creditors should therefore pay more attention to accounting-based information. This result supports H3. Moreover, the positive coefficient of *IRISK* indicates that, for firms with high information asymmetry, one should impose a relatively high weight on accounting-based information. This result is consistent with H4. Healy and Palepu (1993) take the information communication perspective and point out that managers prefer to incorporate as much of the impact from current economic events and private information as possible into current-period financial statements. Managers can leverage the accounting-based information involved in financial statements to reveal their private knowledge to investors. Our results indicate that the effect is particularly pronounced when the level of information asymmetry between market participants and a firm's corporate executives is high.

Lastly, the coefficient of the crisis dummy (i.e., *CRISIS*) is significant and negative. This result supports H5: the loadings on market-based (accounting-based) information should be increased (decreased) when the stock market is experiencing a financial crisis. This result is consistent with the literature on stock market efficiency (e.g., Kraus & Stoll, 1972) where stock prices respond to new information instantaneously, particularly when the condition of a financial crisis is concerned.

Table 5
Estimation results of the hybrid bankruptcy prediction model with dynamic weights (Model 4).

	Coefficient	S.D.	t-value	p-value
Accounting-based approach				
Constant	1.2894	0.1029	12.5260	0.0000
WC/TA	−0.7489	0.1237	−6.0535	0.0000
RE/TA	−2.1124	0.1953	−10.8190	0.0000
EBIT/TA	−22.2290	1.4215	−15.6382	0.0000
MVE/TL	0.0194	0.0047	4.1520	0.0000
SALES/TA	1.3148	0.1962	6.7010	0.0000
Market-based approach				
Constant	−1.8137	0.7153	−2.5357	0.0056
EQAVE	−1.2175	0.2113	−5.7612	0.0000
EQSTD	0.5374	0.2610	2.0588	0.0198
Dynamic Loadings				
Constant	5.8891	0.2683	21.9480	0.0000
SIZE	−0.8758	0.0408	−21.4547	0.0000
VOL	−0.2062	0.0199	−10.3598	0.0000
$\|DA\|$	1.0171	0.0773	13.1624	0.0000
IRISK	0.3570	0.0182	19.6586	0.0000
CRISIS	−0.6711	0.1533	−4.3772	0.0000
Log-likelihood	−9202.6			

This table presents the estimation results of the hybrid bankruptcy prediction model with dynamic weights (see Equations (10)–(13) for the details of the model specifications). See Table 1 for variable definitions and sample descriptions. *CRISIS* is a dummy variable defined by the U.S. GDP growth rate (1 if the GDP growth rate is negative and 0 otherwise).

5.4. Forecasting performance: in-sample test

Our results demonstrate how the two bankruptcy forecasting techniques and the five determinant variables of dynamic loadings are related. In particular, our results indicate that market-based information deserves more attention when it comes to examining large and liquid firms, as well as when the market is crashing. By contrast, more attention should be paid to accounting-based information when assessing a firm with a higher degree of discretionary accruals and information asymmetry. The remaining inquiry examines whether the hybrid model with non-monotonic loadings offers banks better bankruptcy prediction performance. To compare the forecasting performance of the alternative bankruptcy prediction models, two common tests, cumulative accuracy profiles (CAPs) and receiver operating characteristic (ROC) curves, and the corresponding accuracy ratios (*AR*) are detailed below.

First, Fig. 1 shows an example of a CAP plot. To plot the CAP curve, we first rank companies by their default probability estimates, from highest to lowest. Next, we construct a graph with the percentage of all the customers on the *x*-axis and the percentage of all the defaults on the *y*-axis. A good model concentrates the defaulters at the highest default probability estimates and therefore the percentage of all defaulters identified (the *y*-axis) increases quickly as one moves up the sorted sample (along with the *x*-axis). We then use the ideal CAP to compute the accuracy ratio (i.e., the ratio of the area under a model's CAP to the area under the ideal CAP). Fig. 2 presents an example of a ROC curve in which the *y*-axis shows the percentage of true defaults that the model correctly classifies as defaults and the *x*-axis measures the percentage of non-defaults that are mistakenly classified as defaults. The accuracy ratio is comparable to the value of the area under the curve (AUC), since $\text{AR} = 2 \times \text{AUC} - 1$.

Next, Figs. 3 and 4 show, respectively, the CAP and ROC curves of the four bankruptcy prediction models (two for the models involving just one type of information and two for the hybrid models). Our conclusion is clear. The hybrid model with dynamic loadings (i.e., Model 4) generates a larger area under its CAP and ROC curves in comparison with the other models. Next, in Table 6, we summarize the results of the accuracy ratio (*AR*) as a performance measure of bankruptcy prediction. First, examining the two pure bankruptcy prediction models, we find the market-based approach (Model 2) is associated with a higher accuracy ratio in comparison with the accounting-based approach (Model 1). Second, the accuracy ratios of the two hybrid models are higher than those of the two pure models containing just one type of information. This result implies that the adoption of hybrid models that incorporate both the accounting- and market-based approaches can enhance bankruptcy prediction performance.

Although, compared with the pure models in which just one type of information is included, our forecasting tests show the superior performance of the hybrid approach. The next question is whether the hybrid model with dynamic loadings established in this study outperforms the hybrid model with static loadings obtained with the conventional logit method. Returning to Figs. 3 and 4, we find the hybrid model with dynamic loadings outperforms the model with static loadings, with a larger area under its CAP and ROC curves. Moreover, in Table 6, the *AR* value of the hybrid model with dynamic loadings is higher than that of the hybrid model with static loadings.

5.5. Forecasting performance: out-of-sample test

Our results demonstrate that the hybrid model with non-monotonic loadings has better prediction performance with regard to corporate bankruptcy. While the in-sample performance tests of various alternatives give an indication of their historical performance, investors are more concerned with how well they can do in the future using alternative forecasting techniques. We therefore conduct an

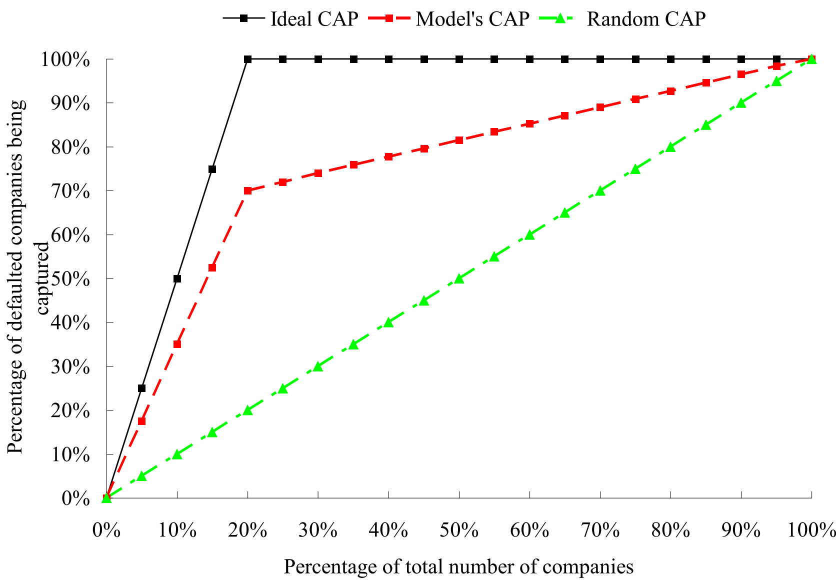


Fig. 1. Illustrations of CAP curves.

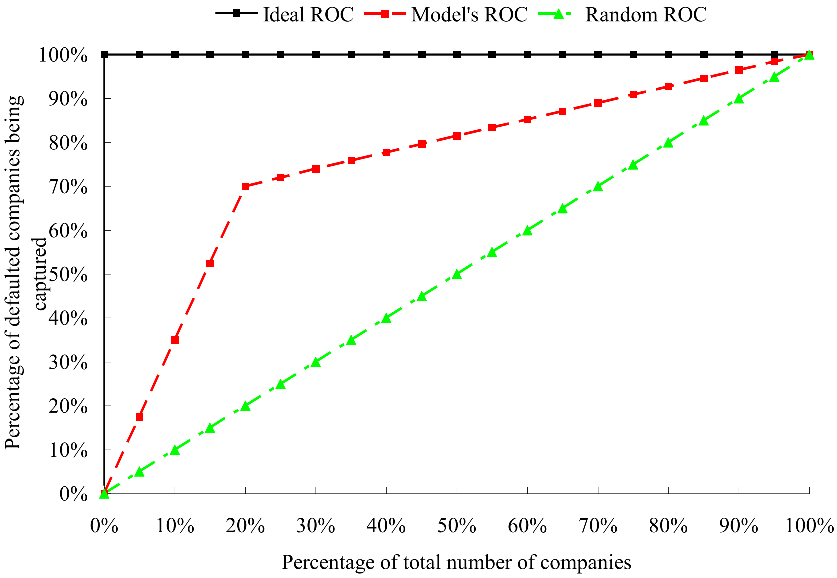


Fig. 2. Illustrations of ROC curves.

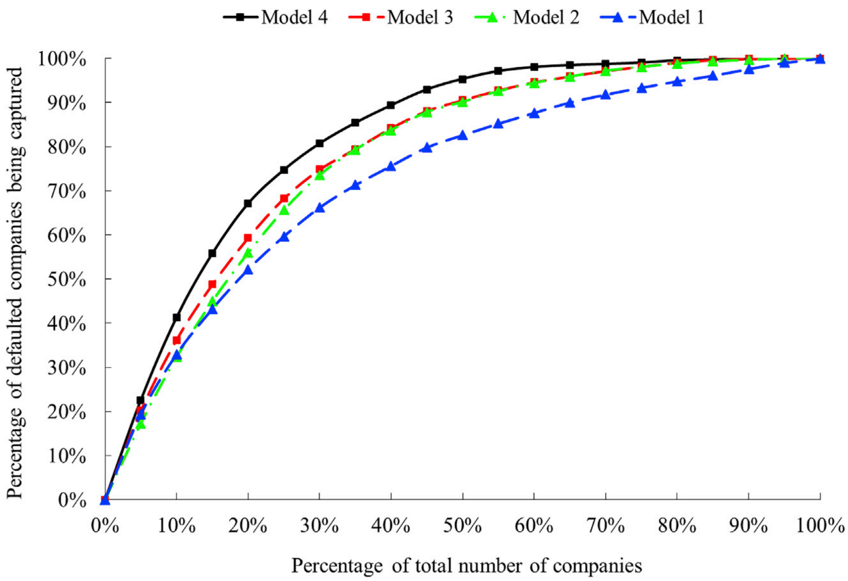


Fig. 3. CAP curves for the alternative bankruptcy prediction models.

out-of-sample test by randomly withholding 10% of the sample, that is, 3114 observations (623 failed and 2491 non-failed firm observations), which are defined as the test set. The residual observations (28,301 = 31,415–3114)) are defined as the model set and are used in calibrating the models and calculating the loadings. The calibrated models are then used to predict the default events in the previously withheld test set.

First, the estimation results of Model 4 (i.e., the hybrid bankruptcy prediction model with dynamic loadings) are listed in Table 7. Importantly, we find qualitatively the same results as those reported in Table 5. In particular, *SIZE, VOL,* and *CRISIS* are significantly negative while $|DA|$ and *IRISK* are significantly positive, supporting our five hypotheses regarding the dynamic weights on the two types of information for bankruptcy prediction. These results also indicate that, while deletion of the observations of the test set slightly changes the values of the parameter estimates, it does not affect their significance or sign. Our previous findings of when creditors should pay more attention to accounting- or market-based information are therefore robust in this subsample. Moreover, these results provide evidence that the multicollinearity issue should not seriously affect our inferences.

Table 8 presents the results of the out-of-sample bankruptcy prediction performance of the alternative bankruptcy prediction models and the corresponding CAP and ROC figures are plotted in Figs. 5 and 6, respectively. Our results are consistent with the following

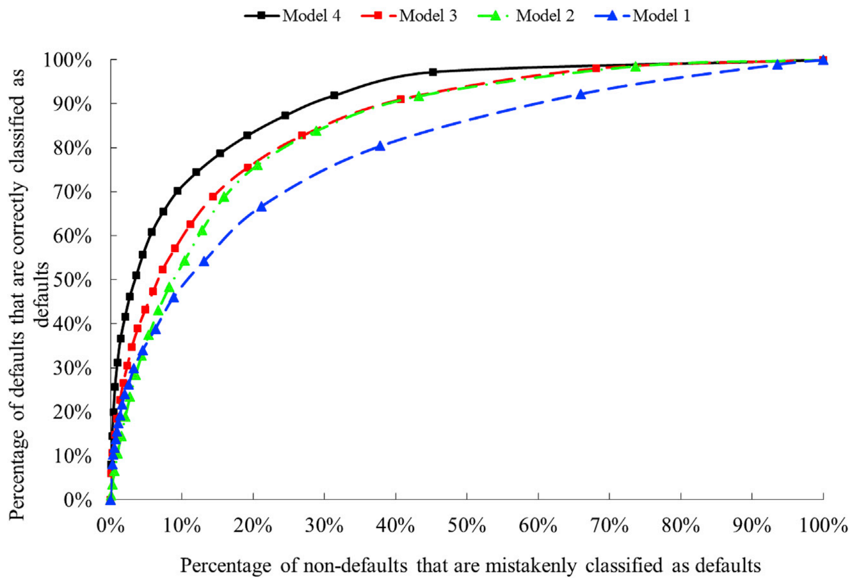


Fig. 4. ROC curves for the alternative bankruptcy prediction models.

Table 6
Accuracy ratios comparison: In-sample tests.

	Accuracy ratio by CAP curve	Accuracy ratio by ROC curve
Highbred bankruptcy prediction models		
Accounting-based approach (Model 1)	56%	58%
Market-based approach (Model 2)	66%	69%
Hybrid bankruptcy prediction models		
Uniform loadings (Model 3)	69%	72%
Non-uniform loadings (Model 4)	**77%**	**80%**

This table summarizes the results of accuracy ratio (*AR*) as a performance measure of bankruptcy prediction. The value in bold denotes the maximum in the column.

Table 7
Estimation results of the hybrid bankruptcy prediction model with dynamic weights (Model 4) using observations from the "model set".

	Coefficient	S.D.	t-value	p-value
Accounting-based approach				
Constant	1.3066	0.1195	10.9307	0.0000
WC/TA	−0.6506	0.1876	−3.4681	0.0003
RE/TA	−2.1918	0.1972	−11.1121	0.0000
EBIT/TA	−21.6343	1.5513	−13.9461	0.0000
MVE/TL	0.0179	0.0039	4.5754	0.0000
SALES/TA	1.2564	0.2080	6.0404	0.0000
Market-based approach				
Constant	−1.6189	0.6021	−2.6889	0.0036
EQAVE	−1.1704	0.1906	−6.1413	0.0000
EQSTD	0.4826	0.2167	2.2270	0.0130
Dynamic Loadings				
Constant	5.9884	0.2719	22.0234	0.0000
SIZE	−0.9082	0.0417	−21.7682	0.0000
VOL	−0.2016	0.0211	−9.5605	0.0000
|DA|	1.1608	0.4022	2.8860	0.0020
IRISK	0.3599	0.0192	18.7438	0.0000
CRISIS	−0.7766	0.1718	−4.5201	0.0000
Log-likelihood	−8246.4			

Table 8
Accuracy ratios comparison: Out-of-sample tests.

	Accuracy ratio by CAP curve	Accuracy ratio by ROC curve
Highbred bankruptcy prediction models		
Accounting-based approach (Model 1)	52%	70%
Market-based approach (Model 2)	64%	66%
Hybrid bankruptcy prediction models		
Uniform loadings (Model 3)	66%	69%
Non-Uniform loadings (Model 4)	**76%**	**79%**

This table summarizes the results of accuracy ratio (*AR*) as a performance measure of bankruptcy prediction. The value in bold denotes the maximum in the column. To conduct the out-of-sample test, 3114 firm-quarter observations (623/2491 of them are failed/non-failed firms) are randomly withheld and defined as the "test set". The residual 28,301 (=31,415–3114) observations are defined as the "model set" and are used in calibrating the models and calculating the loadings.

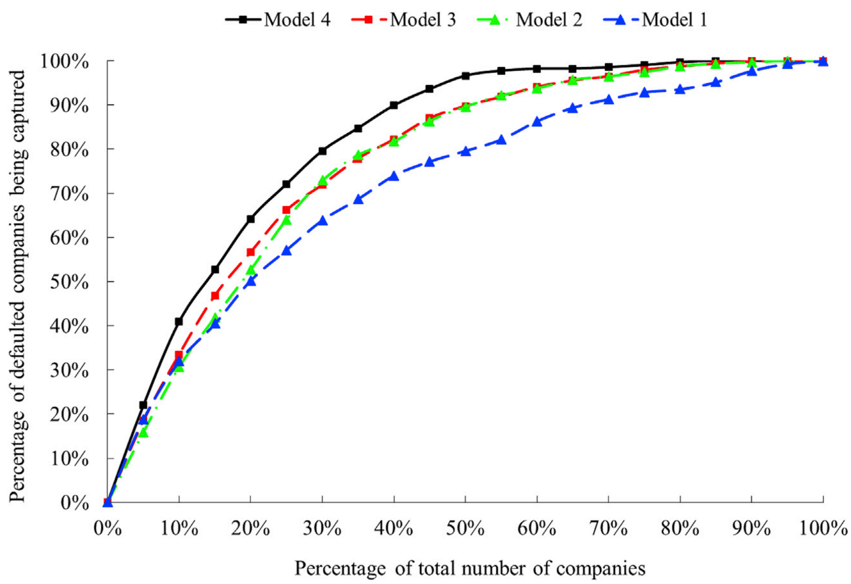


Fig. 5. CAP curves for the alternative bankruptcy prediction models Out-of-sample tests.

notions. First, in having higher *AR* values, the two hybrid models (i.e., Models 3 and 4) still outperform the two pure ones (Models 1 and 2). Second, by having higher *AR* values, the hybrid model with dynamic loadings (Model 4) still outperforms the hybrid model with static loadings (Model 3).

5.6. Dynamic versus statistic weights

A comparison of Equations (9) and (13) clearly reveals the difference between the two hybrid models. Specifically, Equation (9) for the hybrid model with constant weights uses the constant loadings w and 1 - w, to evaluate the impact of each forecasting technique. By contrast, the model specification in Equation (13) employs the endogenous and time-varying loadings $w_{i,t}$ and 1 - $w_{i,t}$. Moreover, Equation (9) could be considered a special case of Equation (13) with the restriction $\theta_1 = \theta_2 = \theta_3 = \ldots = 0$ in Equation (11).

To further compare the two hybrid models, we run Model 4 under the restriction $\theta_1 = \theta_2 = \theta_3 = \ldots = 0$ and the results are shown in Table 9. Comparing these results with Tables 5 and 6, we find a similar value of the log-likelihood function (−10859.90) and the *AR* ratios (*AR* by CAP is 69% and *AR* by ROC is 72%). These results support our argument that Model 3 is a particular case of Model 4 under the restriction of constant weights. Moreover, Table 9 shows that the weights on the accounting- and market-based approaches are 49.91% and 50.10%, respectively, when Model 3 is employed.

Next, Panels A and B of Fig. 7 depict the estimated probabilities of regimes I and II, respectively, which represent the weights for the accounting- and market-based approaches when Model 4 is adopted. First, the estimated weights for each approach flexibly range from zero to one across firm-quarter observations. This finding is consistent with the argument of non-monotonic loadings on accounting- and market-based information. Second, the average values of the estimated probabilities of regime I (i.e., the accounting-based approach) and regime II (i.e., the market-based approach) are 26.41% and 73.59%, respectively. While the former is much smaller than the latter, its 95% confidence interval does not overlap with the values of zero and one. This finding supports the notion that both accounting- and market-based information should be employed when evaluating the credit quality of a company borrower.

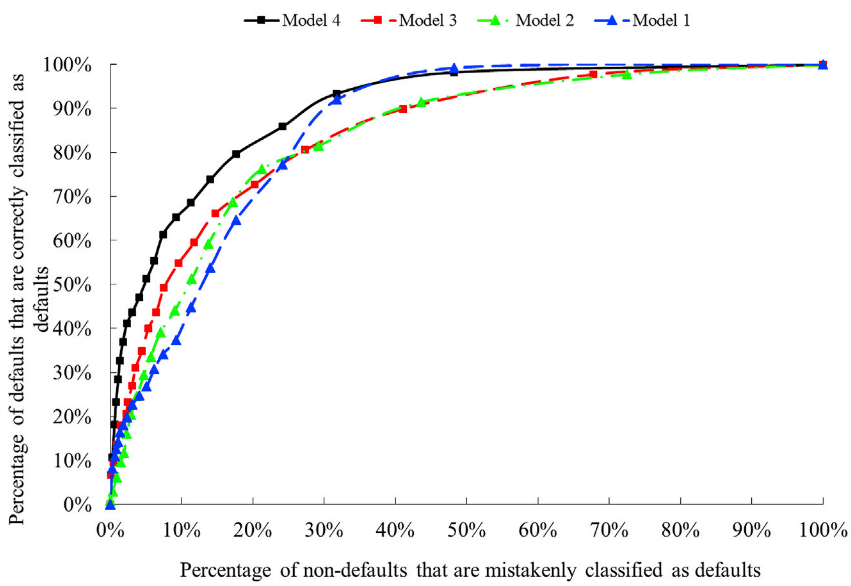


Fig. 6. ROC curves for the alternative bankruptcy prediction models Out-of-sample tests
(A) Accounting-based approach (Regime I)(B) Market-based approach (Regime II).

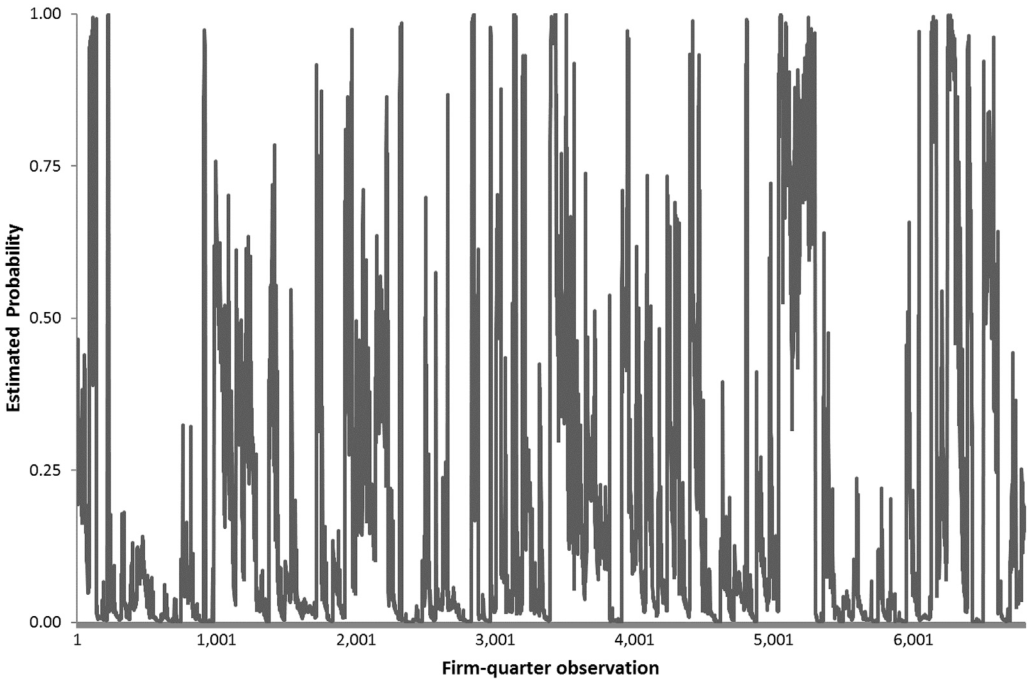
Table 9
Alternative estimation for the hybrid bankruptcy prediction model with static weights (Model 3).

	Coefficient	S.D.	t-value	p-value
Accounting-based approach				
Constant	−3.6177	0.2158	−16.7642	0.0000
WC/TA	0.0162	0.0286	0.5652	0.7140
RE/TA	−0.5052	0.0985	−5.1292	0.0000
EBIT/TA	−56.6568	4.2935	−13.1960	0.0000
MVE/TL	0.1838	0.0163	11.2742	0.0000
SALES/TA	2.8761	0.3513	8.1869	0.0000
Market-based approach				
Constant	7.3141	0.3886	18.8226	0.0000
EQAVE	−2.9779	0.1381	−21.5649	0.0000
EQSTD	0.5918	0.0673	8.7994	0.0000
Transition probability				
Constant (θ_0)	−0.0038	0.0309	−0.1231	0.4510
Log-likelihood	−10859.90			
Weight on ACC	0.4991			
Weight on MKT	0.5010			
AR by CAP	69%			
AR by ROC	72%			

5.7. Explanatory variables for bankruptcy or determinant variables of optimal loadings?

In this study, we propose the five variables to determine the optimal weights on accounting- and market-based information and then develop a hybrid model with dynamic loadings (i.e., Model 4). An alternative model specification is to include all the variables as the explanatory variables for bankruptcy prediction. Therefore, we rerun Model 3 and add all five of the determinant variables of optimal loadings as the explanatory variables in the logit regression. The estimation results are listed in Table 10. Compared with Table 6, the values of the accuracy ratio under the alternative model specification in which all the variables are used as the explanatory variables are equal to those in Model 4 (*AR* by CAP is 77% and *AR* by ROC is 80%). The results are consistent with the notion that one could use these five variables (i.e., *SIZE, VOL,* $|DA|$*, IRISK,* and *CRISIS*) as either the explanatory variables for bankruptcy or the determinant variables of optimal loadings to enhance performance in data fitness.

(A) Accounting-based approach (Regime I)



(B) Market-based approach (Regime II)

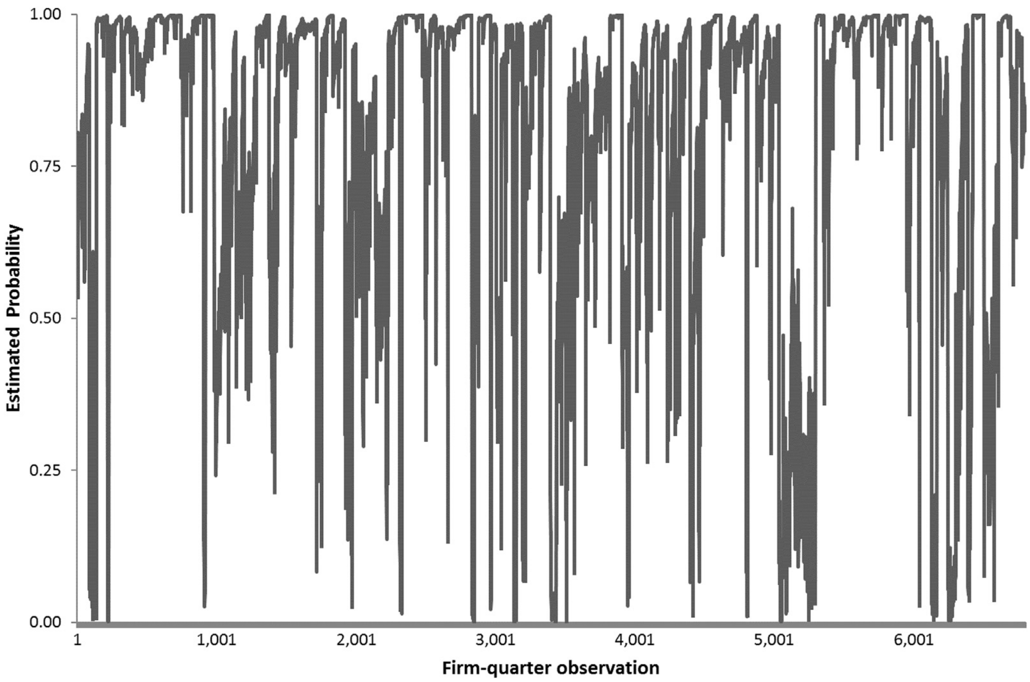


Fig. 7. The estimated weight on accounting- and market-based approach.

6. Conclusions and future research directions

Table 10
Estimation results of the model using all the variables as the explanatory variables for bankruptcy prediction.

	Coefficient	S.D.	t-value	p-value
Constant	5.7694	0.2338	24.6800	0.0000
WC/TA	−1.1576	0.1120	−10.3300	0.0000
RE/TA	−0.2649	0.0414	−6.4000	0.0000
EBIT/TA	−15.5957	0.7345	−21.2300	0.0000
MVE/TL	0.0110	0.0059	1.8600	0.0630
SALES/TA	0.6782	0.0956	7.0900	0.0000
EQAVE	−0.4417	0.0532	−8.3100	0.0000
EQSTD	0.3225	0.0427	7.5500	0.0000
SIZE	−0.7022	0.0234	−29.9900	0.0000
VOL	−0.1460	0.0166	−8.7900	0.0000
|DA|	0.1514	0.2878	0.5300	0.5990
IRISK	10.9945	1.4570	7.5500	0.0000
CRISIS	−0.5554	0.1122	−4.9500	0.0000
Log-likelihood	−9202.4			
AR by CAP	77%			
AR by ROC	80%			

This investigation is one of the first studies to employ a regime-switching technique to establish a hybrid bankruptcy prediction model with endogenous and non-monotonic loadings on accounting- and market-based information. An empirical study based on a data set comprised of default events during the period 1988–2011 is carried out to test the proposed model. In particular, firms that experienced bankruptcy or liquidation events as recorded in the Compustat database are defined as failed firms. Next, for each failed firm, we select three matched non-failed firms drawn from the same industry as the non-failed firm samples.

The empirical results are consistent with the following. First, compared with the models in which just one type of information is involved, in- and out-of-sample tests support the superiority of the hybrid models with regard to bankruptcy prediction performance. Second, the hybrid model with non-monotonic loadings proposed in this study performs better than the hybrid model with constant loadings obtained with the conventional logit method in predicting default events under both in- and out-of-sample settings. Third, to improve the accuracy of bankruptcy prediction, creditors should place more (less) emphasis on market-based (accounting-based) information when large, liquid corporations are considered. Fourth, when predicting the bankruptcy of those companies with incremental information enhanced by accounting discretionary accruals, banks can improve the accuracy by placing more emphasis on accounting ratio-based variables while reducing emphasis on market-based information. Fifth, market participants could increase the loadings of accounting-based information for firms with a high degree of information asymmetry, which is consistent with the notion that managers can reveal their private knowledge to investors using managerial discretion in accounting numbers. Lastly, the loading on market-based (accounting-based) information should be increased (decreased) when the market is crashing, which echoes the market efficiency hypothesis in the literature.

One limitation of this study is that the results are limited by the chosen variables. In particular, we use the five accounting ratio-based variables defined in Altman's Z-score function. Other accounting variables can also be used to explain the bankruptcy events of companies. Since this study's aim is not to find alternative accounting ratio-based variables for bankruptcy prediction, for reasons of expediency, we only use these five common accounting ratio-based variables to establish the accounting-based approach. Next, we obtain the two market-based variables from the conventional DD, in which the value of assets is assumed to equal the value of debt plus equity and the value of debt is supposed to be stable. Different formulations of the DD have been proposed and implemented in the literature. This study considers five variables to control the loading in each accounting- and market-based approach. Future studies could explore other testing variables (e.g., corporate governance addressed by Ali, Liu, & Su, 2018). Bankruptcy prediction has not been satisfactorily modeled. Future studies could also consider other model specifications (e.g., Shumway, 2001). Lastly, banks could encounter corporate borrowers of similar credit quality in practice. Future studies could consider different sample selection designs for non-failed firms.

Data availability

From the sources identified in this paper.

Appendix. Maximum likelihood estimation

To create the best model, we want to find the set of parameters that produces the best fit between the estimated default probability and the observed defaults. In particular, we would like the estimated default probability to be close to 100% for a firm that later defaults and close to zero if the firm does not default. We accomplish this by using maximum likelihood estimation.

In particular, as shown in Equation (13), the estimated default probability for firm i at time t is

$$\begin{aligned} P_{it+1}^{Hybrid} = w_{it+1} \times \frac{1}{1+e^{-(cont_1+\beta_1 WC_{it}/TA_{it}+\beta_2 RE_{it}/TA_{it}+\beta_3 EBIT_{it}/TA_{it}+\beta_4 MVE_{it}/TL_{it}+\beta_t SALES_{it}/TA_{it})}} \\ +(1-w_{it+1}) \times \frac{1}{1+e^{-(cont_2+\gamma_1 MKT\_AVE_{it}+\gamma_1 MKT\_VOL_{it})}} \\ = w_{it+1} \times P_{it+1}^{ACC} + (1-w_{it+1}) \times P_{it+!}^{MKT}, \end{aligned} \tag{A1}$$

where

$$w_{it+1} = prob(s_{it+1}=1|\Pi_{it}) = \frac{\exp(\theta_0+\theta_1\cdot\pi_{1,it}+\theta_2\cdot\pi_{2,it}+\theta_3\cdot\pi_{3,it}+\cdots)}{1+\exp(\theta_0+\theta_1\cdot\pi_{1,it}+\theta_2\cdot\pi_{2,it}+\theta_3\cdot\pi_{3,it}+\cdots)}, \tag{A2}$$

$$1-w_{it+1} = prob(s_{it+1}=2|\Pi_{it}) = 1-prob(s_{it+1}=1|\Pi_{it}). \tag{A3}$$

Next, we define the likelihood function for firm i at time $t+1$ as equal to P_{it+1}^{Hybrid} if the firm does default and $1-P_{it+1}^{Hybrid}$ if it does not:

$$l_{it+1} = P_{it+1}^{Hybrid}, \text{if default}, \tag{A4}$$

$$l_{it+1} = 1-P_{it+1}^{Hybrid}, \text{if no default}, \tag{A5}$$

We then create a log-likelihood function Ψ that is the sum of the natural logarithm of the likelihood functions for all firms:

$$\begin{aligned} \Psi = \sum_{i=1}^{N}\sum_{t=1}^{T} \ln l_{it}(WC/TA_{it}, RE/TA_{it}, EBIT/TA_{it}, MVE/TL_{it}, SALES/TA_{it}, \\ MKT\_MEAN_{it}, MKT\_VOL_{it}; \Omega), \end{aligned} \tag{A6}$$

where Ω is a vector of population parameters containing the unknown elements $cont_1$, β_1, β_2, …, β_5, $cont_2$, γ_1, γ_2, θ_0, θ_1, …. Finally, we use OPTIMUM, a software package from GAUSS, and the built-in Broyden–Fletcher–Goldfarb–Shanno algebra to search for the Ω value that maximizes the above log-likelihood function.[14]

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.iref.2019.02.016.

Panels A and B present the estimation results of the accounting-based and market-based bankruptcy prediction model (see Equations (1)–(6) for the details of the model specifications). See Table 1 for variable definitions and sample descriptions.

To conduct the out-of-sample test, 3114 firm-quarter observations (623/2491 of them are failed/non-failed firms) are randomly withheld and defined as the "test set". The residual 28,301 (=31,415–3114) observations are defined as the "model set" and are used in calibrating the models and calculating the loadings. Other notions are consistent with Table 6.

This table presents the estimation results of the Model 4 under the restriction of $\theta_1=\theta_2=\theta_3=\ldots=0$ in Equation (11), which is equivalent to the Model 3 (i.e., the hybrid bankruptcy prediction model with static weights). See Table 1 for variable definitions and sample descriptions. *CRISIS* is a dummy variable defined by the U.S. GDP growth rate (1 if the GDP growth rate is negative and 0 otherwise).

This table presents the estimation results of the model specification in which all the variables are employed as the explanatory variables for bankruptcy prediction. See Table 1 for variable definitions and sample descriptions. *CRISIS* is a dummy variable defined by the U.S. GDP growth rate (1 if the GDP growth rate is negative and 0 otherwise).

References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance, 32*, 1541–1551.

Ali, S., Liu, B., & Su, J. (2018). Does corporate governance quality affect default risk? The role of growth opportunities and stock liquidity. *International Review of Economics & Finance, 58*, 422–448.

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

Balsam, S. (1998). Discretionary accounting choices and CEO compensation. *Contemporary Accounting Research, 15*, 229–252.

[14] The structural and statistical properties of the OPTIMUM function are well documented in the GAUSS handbook. This study thus omits the statistical properties of the model estimation. For the detail discussion of the bootstrap simulation methods, see Tang, Thomas, Fletcher, Pan, and Marshall (2014) and Lohmann and Ohliger (2017).

Balsam, S., Bartov, E., & Marquardt, C. (2002). Accruals management, investor sophistication and equity valuation: Evidence from 10-Q filings. *Journal of Accounting Research, 40*, 987–1012.
Barry, C., & Brown, S. (1985). Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis, 20*, 407–422.
Bhattacharya, N. (2001). Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review, 76*, 221–244.
Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 7*, 637–654.
Blochlinger, A., & Leippold, M. (2006). Economic benefit of powerful credit scoring. *Journal of Banking & Finance, 30*, 851–873.
Boehmer, E., Jones, C., & Zhang, X. (2008). Which shorts are informed? *The Journal of Finance, 63*, 491–527.
Boehmer, E., & Kelley, E. (2009). Institutional investors and the informational efficiency of stock prices. *Review of Financial Studies, 22*, 3563–3594.
Burgstahler, D. C., & Dichev, I. D. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24*, 99–126.
Burgstahler, D., & Eames, M. (2006). Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting, 33*, 633–652.
Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance, 63*, 2899–2939.
Castagnetti, C., & Rossi, E. (2013). Euro corporate bonds risk factors. *Journal of Applied Econometrics, 28*, 372–391.
Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of Financial Economics, 87*, 249–268.
Chordia, T., Roll, R., & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics, 101*, 243–263.
Chordia, T., Subrahmanyamb, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics, 58*, 41–58.
Das, S. R., Hanouna, P., & Sarin, A. (2009). Accounting-based versus market-based cross-sectional models of CDS spreads. *Journal of Banking & Finance, 33*, 719–730.
Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners and regulators. *Accounting Horizons, 14*, 235–250.
Dierkens, N. (1991). Information asymmetry and equity issues. *Journal of Financial and Quantitative Analysis, 26*, 181–199.
Diether, K., Lee, K., & Werner, I. (2009). Short-sale strategies and return predictability. *Review of Financial Studies, 22*, 575–607.
Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance, 57*, 2113–2141.
Doran, J. S., Peterson, D. R., & Wright, C. (2010). Confidence, opinions of market efficiency, and investment behavior of finance professors. *Journal of Financial Markets, 13*, 174–195.
Hann, R., Lu, Y., & Subramanyam, K. R. (2007). Uniformity versus flexibility: Evidence from pricing of the pension obligation. *The Accounting Review, 82*, 107–137.
Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics, 7*, 85–107.
Healy, P. M., & Palepu, K. G. (1993). The effect of firms' financial disclosure strategies on stock prices. *Accounting Horizons, 7*, 1–11.
Hendershott, T., Jones, C., & Menkveld, A. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance, 66*, 1–33.
Hillegeist, S., Keating, E., Cram, D., & Lundstedt, K. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies, 9*, 5–34.
Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*, 193–228.
Kealhofer, S. (2003). Quantifying credit risk I: Default prediction. *Financial Analysts Journal, 59*, 30–44.
Kealhofer, S., & Kurbat, M. (2001). The default prediction power of the Merton approach, relative to debt ratings and accounting variables. *Moody's KMV*.
Koopman, S. J., & Lucas, A. (2005). Business and default cycles for credit risk. *Journal of Applied Econometrics, 20*, 311–323.
Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics, 39*, 163–197.
Kraus, A., & Stoll, H. R. (1972). Price impact of block trading on the New York stock exchange. *The Journal of Finance, 27*, 569–588.
Krishnan, G. (2003). Audit quality and the pricing of discretionary accrual. *Auditing, 22*, 109–126.
Lee, B. B., & Choi, B. (2002). Company size, auditor type, and earnings management. *Journal of Forensic Accounting, 3*, 27–50.
Li, L., & Miu, P. (2010). A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach. *Journal of Empirical Finance, 17*, 818–833.
Löffler, G. (2007). The complementary nature of ratings and market-based measures of default risk. *Journal of Fixed Income, 17*, 38–47.
Lohmann, C., & Ohliger, T. (2017). Nonlinear relationships in a logistic model of default for a high risk installment portfolio. *The Journal of Credit Risk* (in press).
Maddala, G. S., & Lahiri, K. (2009). *Introduction to econometrics* (4th ed.). The John Wiley and Sons, Ltd., Publication.
Mensah, Y. M. (1984). An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. *Journal of Accounting Research, 22*, 380–395.
Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance, 29*, 449–470.
Miller, R. (1998). *Refmmg ratings, Risk, 11*, 97–99.
Mitchell, J., & Roy, P. V. (2008). Failure prediction models: Performance, disagreements, and internal rating systems. In *European financial management annual conference (athens, Greece)*.
Oderda, G., Dacorogna, M., & Jung, T. (2003). Credit risk models: Do they deliver their promises? *A quantitative assessment, Review of Banking, Finance and Monetary Economics, 32*, 177–195.
Reisz, A., & Perlich, C. (2007). A market-based framework for bankruptcy prediction. *Journal of Financial Stability, 3*, 85–131.
Roll, R., Schwartz, E., & Subrahmanyam, A. (2007). Liquidity and the law of one price: The case of the Futures/Cash Basis. *The Journal of Finance, 62*, 2201–2234.
Sankar, M., & Subramanyam, K. R. (2001). Reporting discretion and private information communication through earnings. *Journal of Accounting Research, 39*, 365–386.
Saunders, A., & Allen, L. (2002). *Credit risk measurement: New approaches to value at risk and other paradigms* (2nd ed.). New York: Wiley Finance.
Schuermann, T. (2005). A review of recent books on credit risk. *Journal of Applied Econometrics, 20*, 123–130.
Scott, R. W. (2000). *Financial accounting theory* (2nd ed.). New Jersey: Prentice Hall.
Shumway, T. (2001). Forecasting bankruptcy more accurately. A simple hazard model. *Journal of Business, 74*, 101–124.
Stein, R. (2005). The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing. *Journal of Banking & Finance, 29*, 1213–1236.
Tang, L., Thomas, L., Fletcher, M., Pan, J., & Marshall, A. (2014). Assessing the impact of derived behaviour information on customer attrition in the financial service industry. *European Journal of Operational Research, 236*, 624–633.
Tian, S., & Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics & Finance, 51*, 510–526.
Tucker, J. W., & Zarowin, P. (2006). Does income smoothing improve earnings informativeness? *The Accounting Review, 81*, 251–270.
Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *The Journal of Finance, 59*, 831–868.
Watkins, J., Vasnev, A. L., & Gerlach, R. (2014). Multiple event incidence and duration analysis for credit data incorporating non-stochastic loan maturity. *Journal of Applied Econometrics, 29*, 627–648.
Wu, J. (2007). *Short selling and the informational efficiency of price, Working Paper*. Texas A&M University.