Performance of default-risk measures: the sample matters

Isabel Abinzano[1], Ana Gonzalez–Urteaga, Luis Muga
INARBE and Public University of Navarre

Santiago Sanchez
Public University of Navarre

Abstract

This paper examines the predictive power of the main default-risk measures used by both academics and practitioners, including accounting measures, market-price-based measures and the credit rating. Given that some measures are unavailable for some firm types, pair wise comparisons are made between the various measures, using same-size samples in every case. The results show the superiority of market-based measures, although their accuracy depends on the prediction horizon and the type of default events considered. Furthermore, examination shows that the effect of within-sample firm characteristics varies across measures. The overall finding is of poorer goodness of fit for accurate default prediction in samples characterised by high book-to-market ratios and/or high asset intangibility, both of which suggest pricing difficulty. In the case of large-firm samples, goodness of fit is in general negatively related to size, possibly because of the "too-big-to-fail" effect.

Key words: credit-risk measures, default prediction, hard to value stocks.

JEL Classification: G32, G33.

1. Introduction

Credit risk is perceived as the oldest and most important form of financial risk. This is because default is one of the most disruptive events that can befall a company, triggering not only bankruptcy costs in the form of legal and consulting fees, but also causing breaks in

[1] Corresponding author. Institute for Advanced Research in Business and Economics (INARBE) and Department of Business Management. Public University of Navarre, Pamplona, Navarre, Spain. E-mail: isabel.abinzano@unavarra.es.
We are grateful for valuable comments from participants at the 11th International Risk Management Conference, the 8th GIKA Conference, the 26th Finance Forum, the 2019 INFINITI Conference on International Finance and the 6th International Conference on Credit Analysis and Risk Management. This paper has been possible thanks to the SANFI Research Grant for Young Researchers Edition 2015, the financial support from the Spanish Ministry of Economy, Industry and Competitiveness (ECO2016-77631-R (AEI/FEDER, UE)) and the Spanish Ministry of Science and Innovation (PID2019-104304GB-I00/AEI/10.13039/501100011033). Ana Gonzalez-Urteaga particularly acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities through grant PGC2018-095072-B-I00.

productivity and the supply chain (Brogaard et al., 2017), in addition to the impact on many parties associated with the defaulting firm (Lyandres and Zhdanov, 2013) and the social costs of economic downturns and recessions (Jones et al., 2017).

In light of this, it is easy to understand the importance of the accurate measurement of default risk. As indicated by Jones and Hensher (2004), Duan et al. (2012) and Tian and Yu (2017) among others, distress forecasts are widely used for a range of purposes, including the solvency monitoring of financial and other institutions by regulators, loan security assessment, going-concern evaluations by auditors, and the measurement of portfolio, credit derivative and other forms of securities-related risk.

Since Beaver's (1966) pioneering work, a great variety of credit-risk measures have been proposed and widely used both by practitioners and academics. The most classic models, such as Altman's (1968) Z-score or Ohlson's (1980) O-score, are based on accounting data. Others use spreads on corporate debt instruments, traditionally bond spreads and, more recently, credit default swap (CDS) spreads. Another alternative is the group of measures based on the price of corporate equity, such as Moody's KMV model or the so-called Black-Scholes-Merton (BSM) measure.[2] A final example would be the assessments of the credit worthiness of firms provided by rating agencies.

Most of the previous literature has shown that market-based measures given by BSM or similar are better default predictors at the one-year horizon. However, most studies do not use alternative market measures, such as bond or CDS spreads, and they simultaneously compare only two types of measures (Kealhofer, 2003; Hillegeist et al., 2004; Gharghori et al., 2006; Hilscher and Wilson, 2017); those that do compare more than two use CDS spreads or ratings rather than actual corporate default data as their reference for checking accuracy (Tanthanongsakkun and Treepongkaruna, 2008; Das et al., 2009; Trujillo-Ponce et al., 2014).

Furthermore, the aforementioned authors compare the predictive power of some credit-risk measures using a single sample, and therefore tend not to pause to consider whether their results might be due to their sample characteristics. However, other studies, such as Shumway (2001), Campbell et al. (2008), Duan et al. (2012) and Li and Faff (2019), instead of using the above measures as a proxy, have established hybrid bankruptcy prediction models incorporating both accounting and market information, and have proved that size, book-to-market (BTM), volatility, liquidity and profitability are key attributes for predicting a company's future default probability.

[2] Also known as Merton's model (see for example Kealhofer, 2003).

Taking into account previous studies, this paper seeks to fill two gaps in the research by analysing the predictive power of the credit-risk measures most commonly used in the literature and examining the influence of sample firm characteristics on their predictive power. The first objective of this paper, therefore, is to evaluate the performance of eight different credit-risk measures using actual information on the occurrence or non-occurrence of corporate credit events as a reference. Specifically, we consider four accounting measures: Altman's Z, Ohlson's O, Zmijewskyi's (1984) model and the probability of Hannan and Hanweck (1988); three market-based measures: CDS spreads, bond spreads and the BSM model; and, finally, the credit rating, which we include in a third category labelled "expert opinion". Additionally, we take account of the fact that due to the large number of variables and the nature of the information involved, data for all of these measures are not available for all firms. For example, not all the companies are rated by a credit rating agency or have a CDS issued on them. As a consequence, the predictive accuracy may be influenced by the availability of data and, ultimately, by the types of companies that make up our sample when analysing each measure of credit risk. Firms in each of the samples will surely have different characteristics (such as size, or intangibility). For this reason, in addition to the non-matched samples, that is, samples composed of the firms with information about a specific default-risk measure, matched samples have also been analysed, that is, samples made up of companies for which data is available for each pair of default risk measures.

Overall, the predictive power of the various measures is seen to increase when considering only severe default events[3], whereas observations across different forecast horizons reveal that it declines over longer horizons for all types of default and for severe default events alone. Furthermore, the three market-based measures provide a better goodness-of-fit than the rest of the measures, with CDSs showing more consistent behaviour in the different scenarios considered. Unlike bond spreads and BSM, CDSs are the only ones that maintain high levels of accuracy regardless of the prediction horizon and the type of default. In addition, it is worth noting the improvement in the accuracy of the credit rating, clearly outperforming the rest when only severe events are considered, even over longer horizons. Finally, the variation obtained in the performance of the measures when distinguishing between non-matched and matched samples suggests that when selecting the measure that best predicts default risk, the sample used and the characteristics of the firms that are part of it are relevant.

[3] By severe default events, we mean those strictly bankruptcy-type events, omitting non-bankruptcy events, such as deferred interest or non-payment of dividends.

Thus, a second objective is to examine the influence of the sample characteristics on the predictive accuracy of the various default risk measures analysed, given that some measures are only available for certain types of firms. In particular, a bootstrap procedure is used to measure the effect of sample characteristics such as size, book-to-market ratio, intangibility and volatility, all linked by the literature to firms that are hard to value or arbitrage and those with higher uncertainty (Baker and Wurgler, 2006; Jiang et al., 2005). In addition, we also analyse the effect on the predictive accuracy of other variables such as liquidity and profitability, which have been shown to be relevant for predicting corporate bankruptcy (Duan et al., 2012), or operating cycle following Chava and Jarrow (2004), who demonstrate the importance of introducing industry effects when assessing the accuracy of default prediction. Among all the variables studied, those most clearly shown to have a major impact on predictive power are size and the BTM ratio. More concretely, we observe that the BTM ratio or intangibility reduce the goodness of fit, suggesting pricing difficulty, whereas volatility enhances it. In contrast, in samples of large firms, predictive capacity is found to be low and decreasing as firm size increases, which seems to be related to the "too-big-to-fail" (TBTF) effect.

In short, the results show that when choosing one measure or another, in addition to the prediction horizon, or even if the intention is to predict all default events or only serious default events, the sample used and the characteristics of the firms that form part of it are relevant. For example, in the case of BSM, which is by far the most widely studied market-based measure in the literature, we observe a clear improvement for the matched sample compared to the non-matched sample, with accuracy values of around 80% and 60%, respectively, when considering all types of defaults and a one-year time horizon. It is therefore evident that the sample matters, as confirmed by the bootstrap analysis. Size and volatility are characteristics that influence the predictability of BSM, which shows high predictive power for small firms and firms with high volatility. For the rest of the measures, the variability observed in accuracy also indicates that the sample matters.

Useful practical implications can be drawn from these findings. The ordering of credit-risk measures in different contexts should be taken into account by investors to make portfolio allocation decisions. It also has important implications in the pension and mutual funds framework when deciding on the target companies to invest in, as well as from a regulatory point of view in determining bank regulatory capital associated with credit risk in the context of the Basel III Standardised Approach.

The rest of this paper is organised as follows. Section 2 reviews the previous literature. Section 3 describes the models and measures of credit risk analysed in the paper. Section 4 presents the database. Section 5 shows the results of the analysis of the predictive power of credit-risk measures. Section 6 analyses the effect of sample characteristics on the different measures of default risk using a bootstrapping procedure. Section 7 presents some robustness checks, and the paper closes in Section 8 with the main conclusions.

2. Literature review

The eight measures mentioned above have been used interchangeably by practitioners and academics to quantify credit risk; however, their credit-risk assessment may vary. For example, in the study of the relationship between credit risk and the momentum effect, several authors, using different measures to proxy for credit risk, obtain different results. Thus, Avramov et al. (2007) use credit ratings, Abinzano et al. (2014) use the BSM model, and Agarwal and Taffler (2008) use the Altman's Z-score, categorised as a binary variable to distinguish between financially-distressed firms and healthy ones. The differences in the results may be due to the different methods used to proxy credit risk.

As a matter of fact, some authors find differences in how closely the various measures are able to reflect the actual levels of the credit risk of firms. Kealhofer (2003) compares the market-based KMV model with credit ratings, and shows that the KMV model is better than credit ratings at predicting and measuring default risk. Hillegeist et al. (2004) compare the BSM probability with the Altman's Z-score and Ohlson's O. Their results show that the BSM probability contains significantly more information about the likelihood of bankruptcy than any of the accounting-based measures and recommend its use as a powerful proxy for bankruptcy probability. In a search for the best-performing measure, Gharghori et al. (2006) compare the BSM model, the BSM model with the equity modelled as a path-dependent barrier option instead of as a standard call option, and an accounting-ratio model similar to the Z-score. The outcome of their analysis indicates that the option-based models clearly outperform the accounting-ratio model as a measure of default risk, and that the performance of the two option-based models is quite similar, leading the author to recommend the simpler BSM model. Hilscher and Wilson (2017), meanwhile, compare the information in corporate credit ratings with that provided by the default prediction model devised by Campbell et al. (2008), which is based on publicly-available accounting and market-based measures. They find that ratings are relatively poor predictors of corporate failure. All these authors reach similar conclusions: the stock market-based BSM method better predicts default at the one-year horizon. It must be stressed, however, that these papers do not analyse alternative market measures such as bonds or CDSs and that, in addition, they simultaneously compare only two types of measures: market-based measures versus the credit rating, market-based measures versus accounting-based measures, or market-based measures amongst themselves.

Admittedly, other authors simultaneously compare more than two types of measures. Thus, Tanthanongsakkun and Treepongkaruna (2008) examine the ability of the BSM model incorporating firm size and book-to-market ratios to explain credit ratings as compared with two accounting ratios, namely, interest coverage and debt leverage ratios, finding the market-based model to be more accurate for this purpose than the accounting-based models. Das et al. (2009), using CDS spreads as a reference, find some accounting ratios to have explanatory power comparable to that of the BSM model's distance to default (DtD). Also using CDS spreads, Trujillo-Ponce et al. (2014) find little difference between the explanatory power of accounting ratios and market-based information given by the DtD and other market-based variables. Although it must be noted that these papers use CDS spreads or ratings rather than actual corporate default data as their reference for checking accuracy, it should also be pointed out that CDS spreads or ratings are themselves credit-risk measures and should therefore not be taken as a reference without first being verified as true credit-risk indicators.

As already stated, one issue to be considered is the influence of sample characteristics on the predictive capacity of the various credit-risk measures. Hilscher and Wilson (2017), for example, find differences in terms of size, leverage and volatility depending on whether the company is rated or not. The literature contains various studies that reveal these variables as important attributes affecting the forward default probabilities of firms. Thus, Shumway (2001) develops a hazard model for forecasting bankruptcy that links market size, stock returns, volatility, leverage and profitability of the company to good out-of-sample accuracy. Campbell et al. (2008) develop a reduced-form model to explore the determinants of corporate failure and find that default is determined by size, book-to-market ratio, leverage, profitability, liquidity, returns and volatility. For their part, Duan et al. (2012) propose a hybrid model combining common factors and firm-specific factors, including the DtD obtained by the BSM model and some accounting ratios. They find that size, book-to-market, volatility, liquidity and profitability have a significant impact on default probabilities. Duan and Miao (2016) later demonstrated the importance of including correlations when trying to obtain an accurate measure of credit risk and to evaluate the possibility of several companies defaulting at once. Li and Faff (2019), who establish another hybrid bankruptcy prediction model based on accounting and market information, conclude that size is a key variable in predicting bankruptcy. They also find that the weight of accounting- versus market-based information varies across companies. Market-based information should carry more weight for large and liquid companies, and accounting-based information should be more influential for those companies characterised by information asymmetry.

For some of these characteristics, BTM in particular, the literature on the topic has reported a more significant level of pricing bias in firms that are hard to value and arbitrage (Baker and Wurgler, 2006), in stocks with information uncertainty (Jiang et al., 2005) and in stocks receiving limited attention (Abody et al., 2010). Thus, companies with high BTM ratios will tend to be harder to value and this will reduce the accuracy of default predictions. Following the same line of argument, but with respect to size, small firms should be harder to value, and therefore higher predictive capacity should be found for large firms. However, in the particular case of the evaluation of bankruptcy risk, this relationship may be altered by the TBTF effect identified by O'Hara and Shaw (1990), whereby big firms cannot be allowed to fail because of the potential adverse impact on the rest of the sector and the economy at large (Moosa, 2010). Indeed, Brewer and Jagtiani (2013) find than banking organisations are willing to pay a premium for mergers that will push them over the TBTF threshold. Thus, the relationship between size and the predictive power of any measure is subject to variation in accordance with whether the sample is composed of large or small firms.

Given this background, this paper complements the earlier literature by simultaneously evaluating the performance of eight credit-risk measures, including accounting-based, market-based and expert opinion measures. An additional feature of this study is that the analysis is performed using actual information on the occurrence or non-occurrence of corporate credit events, rather than a credit-risk proxy, as a reference. Furthermore, we analyse to what extent the prediction horizon or the type of default events affects the predictive capacity of the different credit-risk indicators. Finally, having identified a lack of evidence concerning the possible influence of sample characteristics on the predictive accuracy of alternative default risk measures, we improve the related literature by filling this gap, at least partially. We contribute to the understanding of this issue, placing special emphasis on characteristics associated with pricing difficulty, such as size, BTM, volatility and intangibility, while also considering other variables demonstrated in the literature to be determinants of credit risk, such as profitability, liquidity or industry effects.

3. Measures of credit risk

In this section, we present the eight measures of credit risk that will be analysed for their goodness of fit to real credit risk. As already stated, they include accounting-based, market-based and expert opinion measures. Those in the first category are Altman's Z, Ohlson's O, Zmijewski's model and the probability of Hannan and Hanweck (1988), while the second category includes CDS spreads, bond spreads and the BSM model. The credit rating is considered an expert opinion measure.

Starting with the accounting models, Altman's Z can be considered the classic measure of default risk. Using discriminant analysis, Altman (1968) attempted to predict defaults from five accounting ratios:

- X_1: Working capital/Total assets
- X_2: Retained earnings/Total assets
- X_3: Market value of equity/Book value of total liabilities
- X_4: Sales/Total assets
- X_5: Earnings before interest and taxes/Total assets

The Z-score was calculated with the following expression:

$$Z = 1.2X_1 + 1.4X_2 + 0.6X_3 + 0.999X_4 + 3.3X_5 \tag{1}$$

According to Altman (1968), a Z-Score greater than 3.0 indicates a low probability of default; a score of between 2.7 and 3 signals the need to be alert; a score of 1.8 to 2.7 means a good chance of default; and a score below 1.8 indicates a very high probability of default.

The second accounting-based model used in this study is the O-Score proposed by Ohlson (1980), which, instead of the five variables used for Altman's Z, is obtained from nine variables, including both financial ratios and specific dummies to enhance the predictability of the model:

$$O = -1.32 - 0.407SIZE + 6.03TLTA - 1.43WCTA + 0.0757CLCA - 2.37NITA - 1.83FUTL + 0.285INTWO - 1.72OENEG - 0.521CHIN \tag{2}$$

where:

- SIZE is the log of the ratio of total assets to the GNP price-level index. The index assumes a base value of 100 for 1985.
- TLTA: Total liabilities /Total assets
- WCTA: Working capital/Total assets
- CLCA: Current liabilities /Current assets
- NITA: Net income /Total assets

- FUTL: Cash flows from operations /Total liabilities
- INTWO: One if net income was negative for the last two years, zero otherwise.
- OENEG: One if total liabilities are greater than total assets, zero otherwise.
- CHIN: $(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$, where NI is Net Income.

As we can observe, in contrast to Altman's Z, the higher the O-Score, the higher the risk of default.

Another classic accounting-based method is the model proposed by Zmijewski (1984), determined by probit analysis according to the following equation:

$$X = -4.3 - 4.5X_1 + 5.7X_2 - 0.004X_3 \tag{3}$$

where:

- X_1: Net income/Total assets
- X_2: Total liabilities/Total assets
- X_3: Current assets/Current liabilities

Turning to Hannan and Hanweck (1988), we find a measure of default probability based on a theoretical framework, using three financial variables: capital ratio, expected return on assets and the estimated variance of assets. Thus, default risk is given by the probability of losses exceeding equity:

$$Probability\left(R < -\frac{E}{A}\right) \tag{4}$$

where *R* is the return on assets, and *E/A* is the equity/assets ratio. Based on Tchebysheff's inequality, Hannan and Hanweck (1988) define the probability of default (*DP*) as:

$$DP = Min\left\{1, \left(\frac{\sigma_R}{E(R)+\frac{E}{A}}\right)^2\right\} \tag{5}$$

where σ_R is the standard deviation of the return on assets and *E(R)* the expected return on assets. As noted by Hillegeist et al. (2004) and Trujillo-Ponce et al. (2014), accounting models have been criticised for the historical nature of the information they use as inputs and for not taking into account the volatility of a firm's assets in estimating its default risk. Thus, more recent credit-risk models in the financial literature use data from the capital markets, in which the shares or bonds issued by the companies are traded. In theory, market prices reflect the expectations of investors with regard to a firm's future performance. As a result, these prices contain forward-looking information, which is ideally suited for calculating the probability that a firm will default in the future.

In this way, market prices can be taken directly as measures of credit risk, as has traditionally been the case with bond spreads. Bond spreads are the difference between the corporate bond yield and the risk-free rate. Accordingly, the higher the bond spread, the higher the probability of default of the company. More recently, the empirical literature on credit risk has focused on CDS spreads (e.g. Das et al., 2009; Ericsson et al., 2009; Forte and Peña, 2009). According to Hull et al. (2004), the relationship y-r=s, should therefore hold approximately, where y-r is the corporate bond spread and s is the CDS spread on the company's debt.

An alternative to using the above-mentioned measures of default risk is to construct a measure using the market share prices of firms, as in Moody's KMV model, Vassalou and Xing (2004), Byström et al. (2005) and Byström (2006), to name but a few. These studies start from Merton's (1974) proposal, which is to consider the firm's own equity value as a European call option on its assets and use the Black and Scholes (1973) formula to calculate the equity value.

The likelihood of default is calculated as the probability that the firm's assets will be less than the book value of its liabilities at debt maturity. Assuming the theoretical distribution implied by Merton's model, that is, the normal distribution, the theoretical probability of default is given by the following expression (see Vassalou and Xing, 2004):

$$P_{def,t} = N\left(- \frac{\ln \frac{V_{A,t}}{D_t} + \left(\mu_t - \frac{\sigma_{A,t}^2}{2} \right)(T-t)}{\sigma_{A,t}\sqrt{T-t}} \right) \tag{6}$$

where $V_{A,t}$ is the value of the firm's assets at time t, μ_t is the expected immediate rate of return on $V_{A,t}$, $\sigma_{A,t}$,, is asset return volatility, D_t is the face value of the debt, T is the maturity period, and $N(\cdot)$ is the cumulative probability of the Normal distribution. To find the values of $V_{A,t}$ and $\sigma_{A,t}$, as in Vassalou and Xing (2004), we use an iterative process starting from the market price of the firm's shares.

Finally, we consider the credit rating, which we include in a different category labelled expert opinion, since the credit rating agencies use accounting information along with other factors for their assessment. This measure has the advantage of being simple and easy to understand, but, as is the case with CDS spreads, it must be taken into account that there is no available credit rating for some stocks, especially those of small firms, and that this could result in a sizebiased sample. The accuracy of this measure is also limited by the fact that a firm's credit worthiness can vary significantly before its credit rating is readjusted. Furthermore, it implies that two firms with the same credit rating will also have the same default risk, although substantial differences in default rates may exist within the same bond rating class, as Crosbie and Bohn (2003) have shown.

4. Data

We apply the eight aforementioned measures to companies listed on the New York Stock Exchange (NYSE) for the period January 1986 to January 2016. Banks, finance companies and insurance companies have been excluded from the analysis because of the peculiarities of their capital structure. Stock prices and accounting data are drawn from Thomson Reuters Datastream[4], while the source for actual rating and default events is Moody's Default and Recovery Database.

In keeping with the nature of the study, we use monthly data for the different variables. Following Vassalou and Xing (2004), we avoid problems stemming from reporting delays by not using the book value of accounting variables for the new fiscal year until four months have elapsed.

In line with other studies[5], in the case of the BSM measure, we calculate the book value of debt as short-term debt plus 50% of long-term debt. Furthermore, we need the risk-free rate in order to obtain the implied value of assets. Since we are considering the probability of default in one year, we take the market yield on U.S. Treasury securities at one-year maturity for the whole of the study period.[6]

For CDS spreads, we use the data available in Datastream for 5-year credit default swaps with a modified restructuring clause, according to the ISDA[7] 2003 Credit Derivatives Definitions (revised in 2014).

In the case of bond spreads, in line with Hull et al. (2004), we apply the constraint that the bonds considered must not be puttable, callable, convertible, or reverse convertible. Bonds must not be subordinated or structured and must be single currency. We also applied a time-

[4] Recommendations by Ince and Porter (2006) were adopted for the use of the Datastream data.
[5] See, for example, Crouhy et al. (2000), Vassalou and Xing (2004) and Gharghori et al. (2006).
[6] Although longer horizons are also analysed later in the paper, the main focus is on the findings obtained for a one-year default horizon.
[7] International Swaps and Derivatives Association

to-maturity filter to eliminate bonds with long maturity, and thus enable comparison with the spreads of 5-year CDSs.

The first column in Table 1 shows the number of companies for which we have data on the measures of interest. The differences in the numbers of observations are due to the availability of the variables required to obtain each measure and/or to the type of information used. For example, only certain companies are evaluated by credit rating agencies, and few have issued CDS. The rest of the data in Table 1 are company default data drawn from Moody's Default and Recovery Database. The number of companies with default or non-default data is given, as is that of defaulted companies. This table also gives the main statistics for each measure, for both defaulted and non-defaulted companies. It is worth noting that all these measures indicate lower credit risk for the companies in the non-defaulted sample.

5. Goodness-of-fit analysis of credit-risk indicators

5.1. Methodology

Based on the above, the objective in this section is to analyse the accuracy of credit-risk measures in predicting real default risk. Comparing the performance across different default prediction models is challenging, since they tend to measure slightly different aspects of default events, use different time horizons and may quantify credit risk using different types of inputs.

As in previous works, such as Cantor and Mann (2003), Kealhofer (2003) or Gharghori et al. (2006), we use some metrics developed by Sobehart, Keenan and Stein (2001), namely, cumulative accuracy profile (CAP) plots and accuracy ratios (AR), to ensure that the observed performance can be reasonably expected to represent the behaviour of the model in practice. While CAP plots are a convenient way to visualise model performance, the AR condenses the predictive accuracy of each risk measure into a single statistic for both types of error: Type I (where the model indicates low risk when, in fact, the risk is high) and Type II (where the model assigns a high probability of credit risk when, in fact, the risk is low).

The CAP curve is constructed by plotting the proportion of defaults experienced by firms with the same or higher credit risk against the proportion of all firms with the same or higher risk. The CAP curve is also known as a "power curve" because it shows the effectiveness of a measure for detecting defaults among the population. The further the curve bows toward the northwest corner, the greater the fraction of all defaults that can be experienced by companies with the highest credit risk. The closer the curve is to the southeast corner, the weaker the information content of the credit-risk assessment.

The AR is the ratio of the area between the CAP curve of a given model and the random performance curve (along the diagonal) to the area between the ideal CAP and the random CAP. It is a fraction between minus one and one. Risk measures with ARs close to zero display little advantage over a random assignment of risk scores while those with ARs close to one provide near-perfect predictive power.[8]

5.2. Results

Based on the reasons given above, we combine CAP curves with ARs for each of the eight credit-risk measures, as shown in Figure 1 and Table 2, respectively. To plot CAP curves, we need to label firm-month observations as default or non-default. Following Sobehart et al. (2001), Cantor and Mann (2003) and Gharghori et al. (2006), the firm-month observations for defaulted firms within the prediction horizon of twelve months from the default date are labelled "default", and all other firm-month observations are labelled "non-default". A glance at Figure 1, Panel A, shows CDS spreads to be superior to all the other measures, followed by bond spreads and BSM[9], in that order. These results are in line with those of Table 2, Panel A, where we present the ARs for each measure using the information from all the companies with data on the occurrence or non-occurrence of credit events. The first row shows, with a fit of 91%, that CDS spreads have the highest degree of predictive power for a time horizon of one year, thus clearly outperforming bond spreads (70%) and, even more significantly, the BSM measure (60%). The remaining measures provide very poor fit, with values ranging from 15% for Hannan and Hanweck's probability (HH) to 41% for the credit rating.

So far, all the default events included in the Moody's Default and Recovery Database have been classed under the default heading. However, we must point out that the accounting-based models were estimated to reflect the possibility of companies going bankrupt, which may explain the low AR values previously found for these measures. Thus, in Panel B of both Figure 1 and Table 2, we show the accuracy ratios of the eight measures for severe default events, that is, omitting non-bankruptcy events, such as deferred interest or non-payment of dividends. The increase in goodness of fit can be appreciated in the first row, which shows outstandingly good accuracy values not only for CDS spreads (91%) but also for bond spreads (94%) and BSM

[8] In other words, if all defaults occur for the highest levels of credit risk, the AR would approach one. If all defaults are distributed randomly regardless of the level of credit risk, the AR would be zero. And, if all defaults occur for the lowest levels of credit risk, the AR would approach minus one.
[9] It must be recalled that we are following other works where the BSM measure is computed from the book value of debt taken as short-term debt plus 50% of long-term debt. Thus, we analyse the accuracy of BSM using full long-term debt, finding that it maintains practically the same fit when debt is computed from half the long-term debt alone.

(82%), the last of these showing the most relevant improvement. Meanwhile, the accounting measures continue to exhibit AR values of around 50% or less; hence it can hardly be claimed that their low predictive power is due to the type of credit event considered.

As already stated, this study follows previous research, by considering one-year-head default probability to explore the accuracy of credit-risk models. In fact, as noted by Du Jardin and Severin (2011) and Du Jardin (2015), reviews on financial distress prediction models indicate that these techniques are reliable for estimating only relatively short-horizon default probabilities, and rarely those extending beyond two years. Nevertheless, many studies show that business failure processes can take a number of years, such that symptoms can be traced back more than twelve months.[10] For this reason, we proceed by examining the performance of the measures when forecasting longer-horizon default probabilities. Specifically, in rows 2 to 5 of Table 2, we show the ARs for two- to-five-year default risk horizons, taking into account all types of default events (Panel A) and severe default events (Panel B). In both cases, we observe a rapid deterioration of the accuracy ratios for all measures, with values falling below 50% from the fourth year. CDS and bond spreads (the latter only in the case of severe default) are the exception, with fit values close to 75% and 60%, respectively, even at five years. We can detect this same pattern in both panels in Figure 1 by observing how the CAP curves approach the 45º line as the default forecasting horizon increases, thus implying lower predictive power.

In short, the above results appear to support the proposition that (particularly in the case of CDS, which show high accuracy also at long horizons) market-based measures outperform accounting-based measures and the credit rating. It must be noted, however, that the sample is not the same across all measures, as can be inferred from Table 1. The results could therefore be misinterpreted, since the sample companies differ in terms of size, book-to-market, and other relevant characteristics. To allow for this, we repeat the analysis for the subset of companies for which we have concurrent data on two specific measures, such that the same sample is being analysed.[11] The objective is to see if the sample, and therefore the characteristics of the companies that comprise it, influences the predictability of the different measures in order to observe how each measure will behave against the rest, considering all

[10] See Hambrick and D'Aveni (1988), Laitinen (1991, 1999), Ooghe and de Prijcker (2008), among others. Thus, Altman et al. (2016) study the predictive ability of both financial and non-financial variables over a long horizon of up to ten years for small and medium-sized private enterprises (SMEs), finding several variables that can help analysts to identify early bankruptcy symptoms even five years and longer prior to failure.

[11] The results of the mean difference test between pairs of subsamples confirm significant differences of the characteristics of the companies that belong to the different subsamples, especially with respect to size and intangibility. The results of these tests are omitted to save space, but available from the authors upon request.

the possible sub-samples. Table 3 shows the accuracy ratios for one-year-ahead default probabilities and all types of default events. It can be seen that in individual comparisons with the other measures BSM outperforms them all. However, a more detailed reading of the results reveals a similar level of predictive power in bond spreads and CDS spreads. These three market measures show very high AR values for practically all the matched-sample combinations, providing a clearly superior fit when predicting default events than that obtained with the rest of the measures. Furthermore, given the minimal differences in accuracy ratios between the three market measures, BSM, CDS and bond spreads (with values between 84% and 92%), we cannot claim that any outperforms the rest. Figure 2 confirms this fact. As the CAP curves intersect, none of the measures have greater predictive power for all levels of credit risk.[12] Finally, it is worth noting the high accuracy of credit ratings (AR 91%) in the specific case in which it is matched with the CDS measure. It is clear that sample characteristics are decisive in the accuracy of certain measurements.

One way to obtain a graphic summary of the information provided in Table 3 is to use a stacked bar chart, by summing for every measure the ARs corresponding to the matched samples with the rest of the measures. This allows us to compare the performance of each measure with the remaining measures in the matched samples. In other words, the comparison of the resulting values will show which measures have higher predictive power in different settings when considering firms with different characteristics. This information, which is condensed in Figure 3, confirms that the three market-based measures (CDS, BSM and bonds) perform best, followed by the credit rating, and that the accounting-based measures come last.

If only severe default events are taken into consideration, a clear overall improvement can be seen in the accuracy ratios of matched-samples.[13] These results help to reveal that, in general, credit-risk measures gain in accuracy when only severe default events are considered. Thus, to select the appropriate measure, it must first be determined which type of credit risk we are aiming to detect. In any case, market measures again outperform accounting measures, which remain very low in the accuracy ranking, despite some improvement. It is also worth mentioning the case of credit ratings, which, with ARs exceeding 90% in all cases, lead the race against BSM, bonds and CDS.[14]

[12] It is important to note that, although the accuracy ratio is a good summary measure, not every increase implies an unequivocal improvement in accuracy. Only when the CAP curves do not intersect, will the model with the AR that summarizes the CAP curve further to the northwest quadrant be the best predictor.
[13] Results are omitted for reasons of space but are available upon request.
[14] We have also checked the accuracy of the different measures for the matched samples considering a longer default forecasting horizon, from two to five years, and distinguishing between all default events and severe default events. In both cases, the findings show a significant loss of predictive power in all measures as the default

Finally, Table 4 shows, for each credit-risk measure, the median of the ARs when matched to the rest of the measures. The objective is to summarise the predictive accuracy of each measure into a single value considering different prediction horizons and different types of events. Once again, the greater predictive power of market-based measures is evident. Secondly, the accuracy ratios deteriorate sharply as the default horizon increases. CDS and bond spreads (the latter only in the case of severe default events) are the exception, with fit values close to 70% and 60% respectively, even at five year horizons. However, it is worth noting the improvement in the accuracy of the credit rating when only severe defaults are considered. In this case, the credit rating is the measure that shows the best fit (with medians above 85%) even on the longest prediction horizons.

In short, the main conclusions obtained from the analysis of the results in this section reveal the importance of taking into account various factors, such as the prediction horizon, the type of default and even the characteristics of the companies, before choosing a measure as a proxy of credit risk. In line with previous literature (Kealhofer, 2003; Hillegeist et al., 2004; Gharghori et al., 2006), we observe that BSM offers great predictive power at a one-year time horizon. However, the other two market measures, CDSs and bonds, not previously studied in the literature, show even better levels of accuracy. In fact, if CDS data are available, this would be the most recommended credit risk measure, since it shows the most consistent behaviour across the different scenarios studied. Alternatively, the credit rating is found to be the best predictor of severe default events, but should not be used to predict all types of default. Finally, the findings obtained by distinguishing between non-matched and matched samples suggest that the sample of companies considered, and ultimately the characteristics of the companies, affect the quality of adjustment of the different credit-risk measures. For instance, although our results from the non-matched sample confirm prior evidence that the rating is a poor predictor of default (Kealhofer, 2003; Hilscher and Wilson, 2017), its extraordinarily good accuracy for the case of severe default in the matched sample suggests that the key lies in the sample and, more specifically, in the characteristics of the companies that compose it.

6. Influence of sample characteristics on the accuracy of default-risk measures

Thus far, we have shown that all default-risk measures do not assess credit risk in the same way. The analyses carried out in the previous section allow us to conclude that in general

horizon increases. The only exception is the rating, which presents accuracy ratios greater than 80% when considering severe default events. It should also be noted that, although for CDSs and bonds there is a reduction of accuracy for longer default horizons, the quality of the adjustment remains at fairly high levels. The results are omitted to save space, but available from the authors upon request.

greater accuracy is achieved by CDS spreads, closely followed by BSM and bond spreads for a one-year prediction horizon, all of which are based on market information, and credit ratings if we consider only severe default events. However, since some measures are not available across the whole sample, we cannot categorically conclude which is best.

From Table 5, which shows the average values of firm characteristics for all companies with available data on all credit-risk measures, it is possible to observe, for example, the differences in the market value of equity between companies with data on the BSM measure and those with data on CDS spreads, bond spreads and credit ratings. As already stated, CDS spreads, bond spreads or credit rating data are available only for certain, usually large-size, companies. Indeed, Hilscher and Wilson (2017) point out that rated firms may differ in key ways from non-rated firms, specifically, in size, leverage and volatility - all essential factors for explaining credit risk. Differences can also be seen in other variables, such as BTM or volatility, including lower average BTM (volatility) for CDS, bond spreads and Ohlson's O (CDS and bond spreads). However, differences between market-based, accounting-based and expert opinion measures show no clear pattern of association with intangibility.

We seek to determine whether the accuracy of a measure varies with the characteristics of the companies it is used to assess. For an initial approximation, we divide the sample into quartiles based on size, BTM, volatility and intangibility, and accuracy ratios are calculated for the eight measures of interest in each quartile and goodness of fit with the top and bottom quartiles is then compared. The results are shown in Panel A of Table 6. Next, given that the ARs of three of the measures (CDS, bond spreads and Altman's Z) could not be calculated for the lowest volatility quartile, due to the absence of default events for the firms in that quartile, Panel B in Table 6 shows the sample split at the median of the four characteristics and compares the goodness of fit of the various measures across the two subsamples thus formed. The application of this procedure on all measures and characteristics enhances the robustness of the findings.

An overall negative relationship (higher AR in the lowest quartiles) is found between the variables for size, BTM and intangibility and the predictive power of the various credit-risk measures, while for volatility the relationship varies between positive and negative depending on the credit-risk measure. Two things must be stressed, however. First, note that the relationship must be interpreted in terms of a potential link not between the characteristic and default risk, but between the characteristic and the predictive power of the default-risk measures. Thus, for example, while it is reasonable to expect a positive link between volatility and default risk, the results in Table 6 indicate a mixed relationship between volatility and the predictive power of the various measures. The other point to be stressed is that the observed relationship between sample characteristics and predictive power is only relevant when the credit-risk measure under consideration exhibits reasonably high goodness of fit in either or both of the extreme quartiles. Therefore, when reading the results, it is important to focus on those measures showing high goodness of fit.[15] In this respect, it can be seen that only the market measures, BSM, bonds and in particular CDS, show high AR values for all four characteristics analysed, which is in line with the results obtained in the previous section for the case of the non-matched sample, 1-year ahead default and all types of default events.[16]

A more detailed reading of Table 6 reveals some interesting findings. The CDS spreads show a negative relationship for size and BTM (higher AR in the small-firm and low-BTM groups), the high AR values being exclusive to those two groups. Thus, it is reasonable to conclude that CDS spreads are good predictors only for small firms and those with low BTM. However, their predictive capacity does not appear to be affected by volatility or intangibility, since, regardless of the positive relationship observed with these characteristics, the goodness-of-fit values are high in both groups. BSM, likewise, has good predictive capacity only for small firms and those with high volatility, and this does not appear to be affected by BTM or intangibility. The results for bond spreads are markedly different, their predictive capacity being affected by all four firm characteristics. Bond spreads perform well, with high ARs for small firms and those with low BTM, high volatility or low intangibility. Finally, the credit rating shows good predictive capacity only for low-intangibility firms, and poor goodness of fit overall for all other characteristics and groups.

The performance of CDS spreads and bond spreads in relation to firm size is particularly relevant. While, in terms of goodness of fit, both show high AR values for small firms, they show a surprisingly poor performance for large firms, for which it can be assumed, a priori, that there is more information available, making valuation easier. This result may be due to the TBTF effect (O'Hara and Shaw, 1990). While the various measures may predict a high risk of insolvency, the consequences, both economic and social, of letting a large firm sink are such that it would not be allowed to happen (Type II error).

[15] We consider high goodness-of-fit levels as AR equal to or higher than 70%.

[16] It should be noted that we are focusing here only on the findings obtained for a default horizon of one year and all default events. Recall that accuracy decreases significantly as the default horizon increases, thus detracting from the significance of any observed relationships. However, the relationships exhibit little change for longer horizons. The results, which hold for only severe default events, albeit with a noticeable increase in the accuracy of the credit rating, are available from the authors upon request.

In order to determine whether the above relationships are statistically significant, we estimate various models using a bootstrap procedure with replacement. We generate 1,000 samples of 100 companies for each credit-risk measure and calculate the accuracy ratio for every subsample and the average values of the main characteristics of the companies included in it. Panel A of Table 7 shows the average AR value of the bootstrapped samples. We observe how, also in the bootstrapped samples, the CDS-spreads measure performs best and achieves the highest average accuracy ratio, followed by bond spreads and BSM, while accounting measures show poor goodness-of-fit values. The results of three models incorporating all types of default events over a one-year period[17] are shown in Table 7, Panel B. The first (Model 1) studies the relationship between the accuracy ratio of measure h, with $h=1,2,\dots,8$, AR_h, and the variables $Size_h$, measured as the natural logarithm of the market value of equity, and the book-to-market ratio, BTM_h:

$$AR_h = \alpha_h + \beta_h Size_h + \gamma_h BTM_h + u_h \tag{7}$$

A second model (Model 2) is estimated including equity volatility, $\sigma_{E,h}$, measured as the standard deviation of the past twelve months of stock returns:

$$AR_h = \alpha_h + \beta_h Size_h + \gamma_h BTM_h + \delta_h \sigma_{E,h} + u_h \tag{8}$$

Finally, Model 3 includes, in addition to size and the BTM ratio, the intangibility ratio, *INTAN*, measured as the ratio of intangible to total assets:[18]

$$AR_h = \alpha_h + \beta_h Size_h + \gamma_h BTM_h + \delta_h INTAN_h + u_h \tag{9}$$

The results displayed in Table 7 are consistent with the previous evidence shown in Table 6. Here again, the sole focus is on analysing the relationship between the firm characteristics and the predictive capacity of the market-based credit-risk models and the credit rating, which are those that provide reasonably high goodness of fit.[19]

First, it can be seen that size has a negative and significant impact on accuracy in the three models estimated for BSM, CDS spreads and bond spreads. That is, the larger the company,

[17] Similar results from the cases incorporating severe default events only and a five-year horizon are available from the authors upon request.
[18] Since the joint inclusion of *BTM*, *INTAN* and σ in the same model triggered multicollinearity issues, the results are not shown.
[19] Thus, in order to save space, we omit the results related for accounting measures. Although some significant relationships are detected for some of them, it is important to recall that the detection of a certain characteristic's positive/negative effect on a certain credit-risk measure is of no relevance if it has poor fit to the data across all levels of that characteristic.

the poorer the model fit (AR).[20] It must be noted that, as seen in Table 5, the companies with available data on these measures are the largest, so the negative sign of the size coefficient might be explained by the so-called TBTF effect, confirming the results previously shown in Table 6. Secondly, and again in line with the results shown in Table 6, the coefficient on the BTM variable is negative and significant for BSM and CDS spreads, and while it is positive for bond spreads, it lacks significance in Model 3. Thirdly, in Panel B, volatility can be seen to have a positive effect on the accuracy ratio of the BSM, CDS spreads and bond spreads. This result may be capturing one of the recognised advantages of market measures; namely, that they take into account the volatility of equity. Finally, the intangibility coefficient is negative and significant in the case of bond spreads, which is consistent with the fact that the more intangible the assets of a company, the harder it is to value, and consequently, the more difficult it is to assess its credit risk.

7. Robustness Checks

7.1. Bootstrapped sample restricted to companies with data for CDS spreads

Despite the above results, however, we must remember that these samples contain companies differing in size, BTM, and other characteristics, as seen in Table 5. In fact, in Table 8, which shows the average values of the characteristics of the bootstrapped samples used in the previous section, it can be observed that the companies with data for CDS and bond spreads are larger than those in the other samples. Therefore, in Table 9 we repeat the estimation with the requirement that companies for measure *h* must also have data on CDS; that is, the measure with the best accuracy ratio for the non-matched sample must also belong to the highest market-value category. After applying this requirement, the results are consistent with those shown previously in Table 7, Panel B. The *Size* coefficient remains negative for BSM, bond spreads and CDS spreads, a result already acknowledged as possibly attributable to the TBTF effect. Similar results are obtained for severe default events and longer default horizons. Notice, however, that even for bigger companies, in the case of the credit rating, the coefficient on size is positive and significant in every model. This result could be due to agencies applying a TBTF correction factor on market information in their estimations.

7.2. Bootstrapped sample analysis including additional characteristics

Previous sections have revealed heterogeneity of the predictive power of credit-risk measures linked to firm characteristics, such as size, BTM, volatility or intangibility. Similarly, Duan et

[20] Size, in contrast, shows a positive and significant effect in the case of the rating, although its AR values are very low across all size levels (see Table 6).

al. (2012) show that variables such as liquidity (cash to total assets) and profitability (net income to total assets) can enhance the predictive power of credit-risk measures. It is therefore reasonable to ask whether the inclusion of these two characteristics in our estimations might show them to be associated with default-risk forecasting accuracy, and whether the relationship varies across the various measures.

Another factor potentially affecting the probability of default is the industry sector. However, it is difficult to obtain a sectoral classification of firms with operations in several different sectors of activity, a problem that becomes especially relevant in the case of relatively large firms. One alternative is to use the operating cycle, which can be inherently linked to the industry sector in which the firm operates (Dechow 1994; Charitou 1997).

To analyse the impact on the AR of the various measures, the previously described bootstrap procedures are repeated while adding, in turn, each of the three variables mentioned, to Model 1, equation (7). The results for the whole sample, 1-year-ahead default and all types of default event are given in Table 10. Panel A shows the goodness of fit of the measures to the top and bottom quartiles as a function of these variables, while Panel B displays the model estimates. In all three of the proposed models, the relationships with size and BTM appear consistent with those shown in Table 7.

We now look at the two new variables adopted from the work of Duan et al. (2012). Initially, we observe a positive and significant effect of liquidity on the accuracy of BSM and CDS spreads (Model 4), although all the market measures (including the rating) show high goodness-of-fit values regardless of the level of liquidity (Panel A). Furthermore, there seems to be no significant relationship between the ARs of the market measures and the profitability of the company (Model 5).[21] Finally, with the operating cycle as the sector proxy (Model 6), it can be seen that longer average periods are associated with higher AR for CDS spreads, although the fact is that the AR values are high regardless of the operating cycle, not only for CDS spreads, but also for BSM and, to a lesser degree, bond spreads.

Thus, despite initial signs of some relationships with these new characteristics, none of them is seen to affect the predictive capacity of the credit-risk measures under consideration. In short, most of the heterogeneity in the predictive power of credit-risk measures that is attributable to firm characteristics appears to be captured by the size and BTM variables, as shown in the

[21] A negative and significant relationship is observed only for CDS. However, a closer look at Table 10, Panel A, reveals that there are no AR data for Q4 (since among the firms with CDS data none of those with high profitability values have default data). Therefore the observed negative relationship is merely an indication that the AR decreases with higher profitability, but only in a low-profitability context.

previous analyses. The main conclusions to be drawn from Sections 6 and 7 are summarised in Table 11.[22]

7.3 Joint Analysis of all measures

Finally, to complement the individual analysis, Table 12 shows the results of a joint analysis of all the credit-risk measures using dummy variables and considering all default events occurring within a one-year horizon. The hidden dummy is the CDS spread, because it has the highest accuracy ratio for the non-matched sample.

Note that a negative sign on the coefficient associated with the dummy variable for any credit-risk measure indicates that it has poorer goodness of fit than CDS spreads. Therefore, given the negative coefficients of all the dummies, we can confirm the superiority of CDS spreads over the rest of the measures. Furthermore, taking into account that the more negative the sign on the coefficient of the dummy, the poorer the accuracy ratio, the results appear consistent with those shown previously, and we are therefore able to confirm the greater predictive power of CDS spreads, bond spreads and BSM measures. The results also indicate that the accuracy ratio is affected positively by volatility, liquidity and operating cycle, and negatively by BTM, intangibility, and size (significantly so in five of the six models considered), while company profitability is not statistically significant.

8. Conclusions

The measurement of credit risk has long been a predominant concern for both academics and practitioners. Credit risk is widely used for a range of purposes and its prediction is crucial to the avoidance of major economic and social consequences. However, relatively little is known about the respective predictive capacities of the most widely used measures. The main aim of this paper is, specifically, to explore and explain the strengths and weaknesses in terms of predictive accuracy of the eight measures most widely used for assessing real default risk in the literature on credit risk: four accounting-based models, two market-data measures, one market-based constructed model and one expert opinion measure.

To evaluate the aforementioned eight measures, we study their discriminatory power to predict all types of default events and severe default events over short and long forecasting horizons and explore the effect of some firm characteristics on their forecasting accuracy. Overall, all measures perform better for severe default events and shorter forecasting horizons.

[22] This table also includes the findings obtained for the cases of severe default events and longer default horizons, which are not included in the main text, but are available from the authors upon request.

Furthermore, the market-based models clearly outperform the accounting-ratio models and the credit rating in terms of outcome prediction. Among the market-based models, CDS spreads show the most consistent behaviour, achieving AR values similar to the others in the prediction of all default types and severe default events within a one-year period, but being the only one that maintains a good fit even over longer forecasting horizons. It also highlights the extraordinary performance obtained in the case of the credit rating by predicting severe default events regardless of the prediction horizon.

In all cases considered, the model fit is found to depend on the sample. Given that data on certain types of credit-risk measures, such as CDS spreads, bond spreads or credit ratings, are available only for large companies, this issue has been carefully studied, since the performance of the models could vary with company characteristics. This study has focused on those measures seen to have high predictive capacity. The results are enlightening, showing that variables such as the book-to-market ratio, or the company's intangibility in general, two proxies for assets that are hard to value or difficult to arbitrage, reduce the goodness of fit of the models, whereas volatility enhances it.

Firm size is also shown to have a negative impact on the predictive capacity of the models. In other words, in large-firm samples, predictive capacity is found to be low and decreasing as firm size increases. This would support the hypothesis that the measures are unable, overall, to capture the TBTF effect and could be indicating high levels of insolvency risk when the large size of a firm makes default unlikely. The exception in this respect would be the credit rating, which, despite not having a high level of predictive capacity for either firm-size category, relates positively with better predictive capacity as firm size increases. This could have something to do with the discretionary adjustment made by rating agencies enabling them to capture the TBTF effect in their ratings.

These results prove robust to the inclusion of other variables, such as liquidity and profitability, which have proved useful in insolvency-forecasting models (Campbell et al., 2008; Duan et al. 2012) and to the use of the operating cycle to capture the firm's industry.

This study serves to confirm that, when trying to identify the most accurate predictors of bankruptcy risk, the sample matters. To avoid misleading conclusions, therefore, any use of possible new bankruptcy risk measures and any comparison of these with existing ones should take into account the characteristics of the sample from which the data are drawn.

Finally, this study holds several practical implications for agents involved in investment decisions. Investors should, where possible, take into consideration market-based credit-risk measures to make portfolio allocation decisions, since these offer the truest reflection of actual default risk. Of all market-based measures, BSM would be the most recommendable both in terms of accuracy and availability, although for samples that do not have the CDS data limitation, CDS is the measure that performs best regardless of the prediction horizon and type of default. The ranking of measures obtained should be also considered in the pension and mutual funds framework, where the credit rating is used as an investment screen. Nevertheless, when using the rating it is crucial to keep in mind that it is a good predictor of severe default events, but its performance is quite poor if all types of default events are considered. Furthermore, not all companies are credit rated, which restricts the number of companies in the portfolio. In addition, this study also has important implications for the regulatory perspective. In the Basel III Standardised Approach, the credit rating is used to determine bank regulatory capital for credit risk. Again, availability and accuracy are two aspects to take into account. Indeed, note that the weight for the unrated corporations is 100%. Furthermore, in order to determine the credit quality of bank debtors, it is important to appropriately predict not only severe defaults, but all types of default, since less severe defaults could also determine the failure of debtors to meet their contractual payment obligations with the bank. However, as shown, the rating fails at this point. An alternative could be the BSM model, due to the broader availability and its high accuracy, at least at short horizons.

References

Abinzano, I., Muga, L., Santamaría, R., 2014, "Is default risk the hidden factor in momentum returns? Some empirical results", *Accounting and Finance*, 54 (3), 671-698.

Abody, D., Lehavy, R., Trueman, B., 2010, "Limited attention and the earnings announcement returns of past stock winners", *Review of Accounting Studies*, 15, 317 – 344.

Agarwal, V. and Taffler, R., 2008, "Does financial distress risk drive the momentum anomaly?", *Financial Management*, 37: 461 – 484.

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

Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K., Suvas, A., 2016, "Financial and Non-Financial Variables as Long-Horizon Predictors of Bankruptcy", *Journal of Credit Risk*¸12, 4.

Avramov, D., Chordia, T., Jostova, G., Philipov, A., 2007, "Momentum and credit rating" *Journal of Finance*, 62 (5), 2503 - 2520.

Baker, M., Wurgler, J., 2006, "Investor Sentiment and the Cross-Section of Stock Returns", *Journal of Finance*, 61, 1645 – 1680.

Beaver, W., 1966, "Financial Ratios as Predictors of Bankruptcy", *Journal of Accounting Research*, 6, 71 – 102.

Black, F., Scholes, M., 1973, "The pricing of options and corporate liabilities", *Journal of Political Economy*, 81, 3, 637 – 654.

Brewer, E., Jagtiani, J., 2013, "How Much Did Banks Pay to Become Too-Big-To-Fail and to Become Systemically Important?", *Journal of Financial Services Research*, 43, 1-35.

Brogaard, J., Li, D., & Xia, Y. (2017). Stock liquidity and default risk. *Journal of Financial Economics*, 124, 3, 486-502.

Byström, H., L. Worasinchai, and S. Chongsithipol, 2005, "Default risk, systematic risk and Thai firms before, during and after the Asian crisis", *Research in International Business and Finance*, 19, 95-110.

Byström, H., 2006, "Merton unraveled: A flexible way of modeling default risk", *Journal of Alternative Investments*, 8, 39-47.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008, "In search of distress risk", *Journal of Finance*, 63, 6, 2899 – 2939.

Cantor, R., Mann, C., 2003, “Measuring the Performance of Corporate Bond Ratings”. *Moody’s Special Comment*. Available at SSRN: https://ssrn.com/abstract=996025.

Charitou, A., 1997, “The role of cash flows and accruals in explaining security returns: evidence of the UK”. *European Accounting Review*, 6, 4, 629–652.

Chava, S., Jarrow, R., 2004, “Bankruptcy Prediction with Industry Effects”, *Review of Finance*, 8, 537-569.

Crosbie, P., Bohn, J. (2003), “Modeling default risk”. *Moody’s KMV*.

Crouhy, M., Galai, D., Mark, R., 2000, “A Comparative Analysis of Current Credit Risk Models”, *Journal of Banking and Finance,* 24, 59-117.

Das, S., Hanouna, P., Sarin A., 2009, “Accounting-based versus market-based cross-sectional models of CDS spreads”, *Journal of Banking and Finance,* 33, 719-730.

Dechow, P.M., 1994, “Accounting earnings and cash-flows as measures of firm performance. The role of accounting accruals”, *Journal of Accounting and Economics*, 18, 3–42.

Duan, J-C., Sun, J., Wang, T., 2012, “Multiperiod corporate default prediction – A forward intensity approach”, *Journal of Econometrics,* 170, 191-209.

Duan, J-C., Miao, W., 2016, “Default correlations and large-portfolio credit analysis*”, Journal of Business & Economic Statistics*, 34, 536–546.

Du Jardin, P., Severin, E., 2011, “Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model”, *Decision Support Systems,* 51, 701-711.

Du Jardin, P., 2015, “Bankruptcy prediction using terminal failure processes”, *European Journal of Operational Research,* 221, 378-396.

Ericsson, J., Jacobs, K., Oviedo-Helfenberger, R, 2009, “The determinants of credit default swap premia”, *Journal of Financial and Quantitative Analysis,* 44, 109-132.

Forte, S., Peña, J.I, 2009, “Credit spreads: An empirical analysis on the informational content of stocks, bonds, and CDS”, *Journal of Banking and Finance,* 33, 2013-2025.

Gharghori, P, Chan, H., Faff, R., 2006, “Investigating the performance of alternative default-risk models: Option-based versus accounting-based approaches”, *Australian Journal of Management*, 31, 2, 207 – 234.

Hambrick, D.C., D'Aveni, R.A., 1988, "Large Corporate Failures as Downward Spirals", *Administrative Science Quarterly*, 33, 1, 1 – 23.

Hannan, T., Hanweck, G.A., 1988, "Bank Insolvency Risk and the Market for Large Certificates of Deposit", *Journal of Money, Credit and Banking*, 20, 2, 203 – 211.

Hillegeist, S.A., Keating, E.K., Cram, D.P., Lundsted, K.G., 2004, "Assessing the Probability of Bankruptcy", *Review of Accounting Studies*, 9, 5-34.

Hilscher, J., Wilson, M., 2017, "Credit ratings and credit risk: Is one measure enough?" *Management Science*, 63, 10, 3414-3437.

Hull, J.C., Predescu, M., White, A., 2004, "The relationship between credit default swap spreads, bond yields, and credit rating announcements", *Journal of Banking and Finance*, 28, 2789–2811.

Ince, O.S., Porter, R. B., 2006, "Individual equity return data from Thomson Datastream: Handle with care!", *Journal of Financial Research*, 29, 463-479.

Jiang, G., Lee, C., Lee, M. C., Zhang, Y., 2005, Information Uncertainty and Expected Stock Returns, *Review of Accounting Studies*, 10, 185 – 221.

Jones, S., Hensher, D.A., 2004, "Predicting Firm Financial Distress: A Mixed Logit Model", *Accounting Review*, 79, 4, 2011-1038.

Jones, S., Johnstone, D., Wilson, R., 2017, "Predicting Corporate Bankruptcy: An Evaluation of Alternative Statistical Frameworks", *Journal of Business Finance and Accounting*, 44 (1-2), 3-34.

Kealhofer, S., 2003, "Quantifying Credit Risk I: Default Prediction", *Financial Analysts Journal*, 59, 1, 30 – 44.

Laitinen, E.K., 1991, "Financial ratios and different failure processes", *Journal of Business Finance and Accounting*, 18, 649 – 674.

Laitinen, E.K., 1999, "Predicting a corporate credit analysis risk estimate by logistic and linear models", *International Review of Financial Analysis*, 8, 2, 97 – 121.

Li, L., Faff, R., 2019, "Predicting corporate bankruptcy: What matters?", *International Review of Economics and Finance*, 62, 1-19.

Lyandres, E., Zhdanov, A., 2013, "Investment opportunities and bankruptcy prediction", *Journal of Financial Markets*, 16, 439-476.

Merton, R.C., 1974, “On the pricing of Corporate debt: The risk structure of interest rates”, *Journal of Finance*, 29, 449 – 470.

Moosa, I., 2010, “The myth of too big to fail”, *Journal of Banking Regulation,* 11, 4, 319-333.

O’Hara, M., Shaw, W., 1990, “Deposit Insurance and Wealth Effects: The Value of Being “Too Big to Fail””, *Journal of Finance,* 45, 1587-1600.

Ohlson, J.A., 1980, “Financial Ratios and the Probabilistic Prediction of Bankruptcy”, *Journal of Accounting Research*, 18, 1, 109 – 131.

Ooghe, H., De Prijcker, S., 2008, “Failure processes and causes of company bankruptcy: a typology”, *Management Decision*, 46, 2, 223 – 242.

Shumway, T., 2001, “Forecasting Bankruptcy More Accurately: A Simple Hazard Model”, *Journal of Business*, 74, 101-124.

Sobehart, J., Keenan, S., Stein, R., 2001, “Benchmarking Quantitative Default Risk Models: A Validation Methodology”, *Algo Research Quarterly,* 4, 1-2, 57-72.

Tanthanongsakkun, S., Treepongkaruna, S., 2008, “Explaining credit ratings of Australian companies. An application of the Merton model”, *Australian Journal of Management,* 33, 2, 261 – 275.

Tian, S., Yu, Y., 2017, “Financial ratios and bankruptcy predictions: An international evidence”, *International Review of Economics and Finance*, 51, 510-526.

Trujillo-Ponce, A., Samaniego-Medina, R., Cardone-Riportella, C. (2014), “Examining what best explains corporate credit risk: Accounting-based versus market-based models”, *Journal of Business Economics and Management,* 15, 2, 253-276.

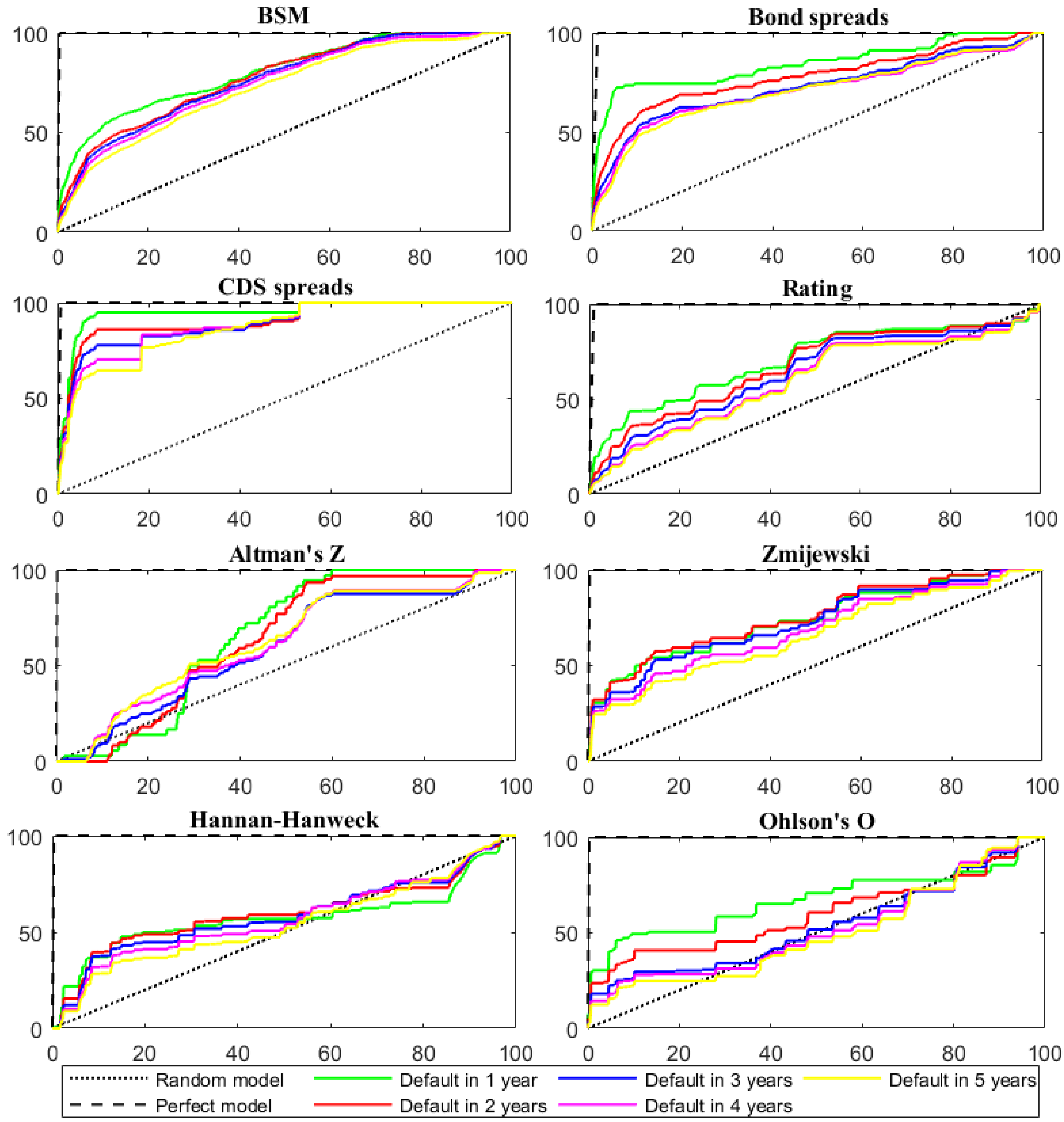
Vassalou, M., Xing, Y., 2004, “Default risk in equity returns”, *Journal of Finance,* 59, 2, 831-868.

Zmijewski, M.E., 1984, “Methodological Issues Related to the Estimation of Financial Distress Prediction Models”, *Journal of Accounting Research,* 22, 59 – 82.

Figure 1: CAP curves for the different measures of credit risk. Non-matched samples.

This figure shows the cumulative accuracy profile (CAP) curves for every measure of credit risk for the non-matched samples considering default forecasting horizons from one to five years and taking into account all default events (Panel A) and only severe default events (Panel B). The CAP curve is constructed by plotting the proportion of defaults against the proportion of firms ordered from highest to lowest credit risk. The further the curve bows toward the northwest corner, the greater the fraction of all default probability assigned to high-credit-risk companies. The closer the curve is to the southeast corner, the weaker the information content of the credit-risk assessment.

Panel A: All default events



Panel B: Only severe default events

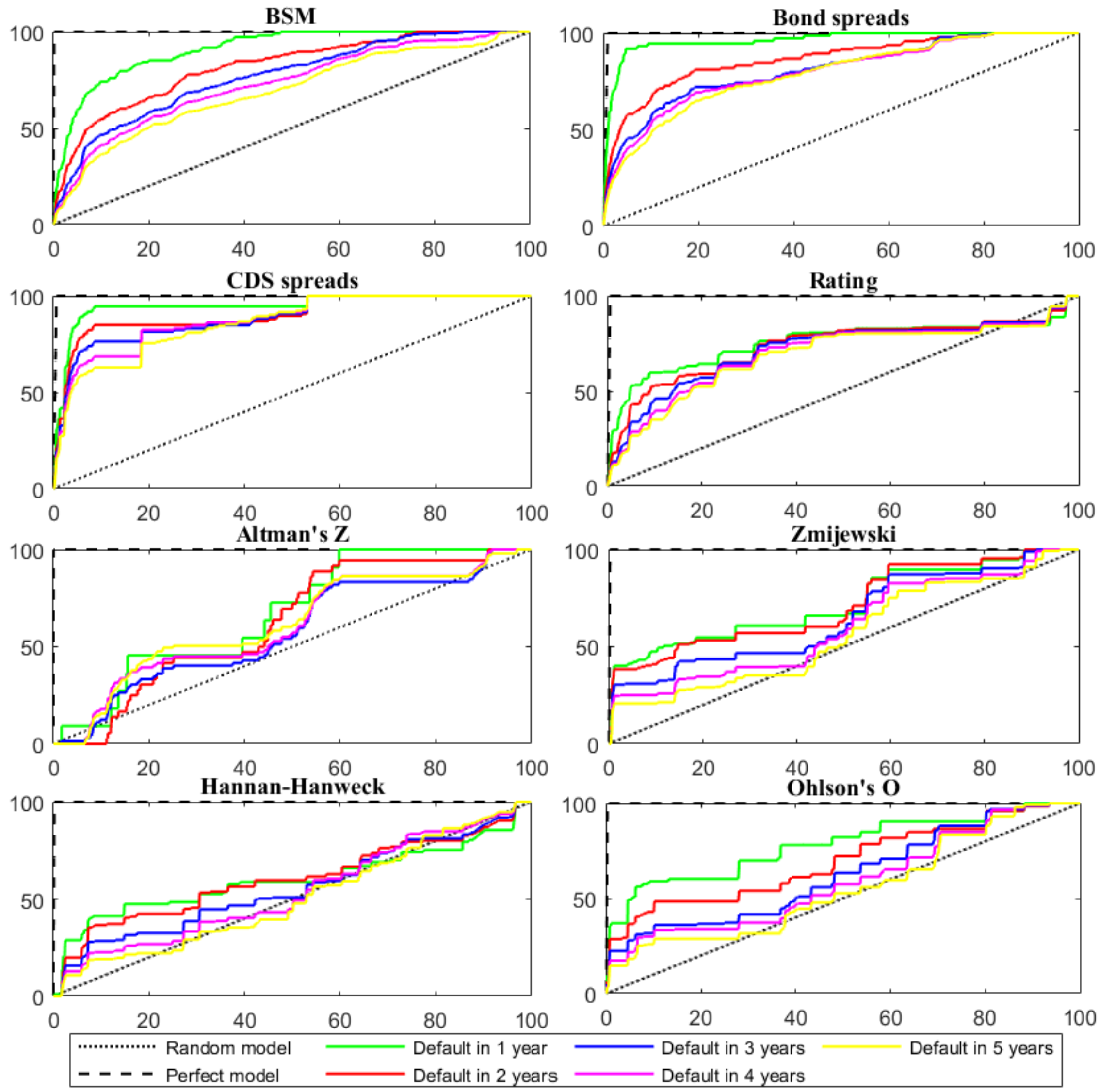


Figure 2: Comparison of CAP curves for the BSM, CDS spreads, bond spreads and credit rating. Matched samples. Default in one year. All default events.

This figure shows the cumulative accuracy profile (CAP) curves for pairs of credit-risk measures for matched samples considering default forecasting horizons of one year and taking into account all default events. The CAP curve is constructed by plotting the proportion of defaults against the proportion of firms ordered from highest to lowest credit risk. The further the curve bows toward the northwest corner, the greater the fraction of all default probability assigned to high-credit-risk companies. The closer the curve is to the southeast corner, the weaker the information content of the credit-risk assessment.

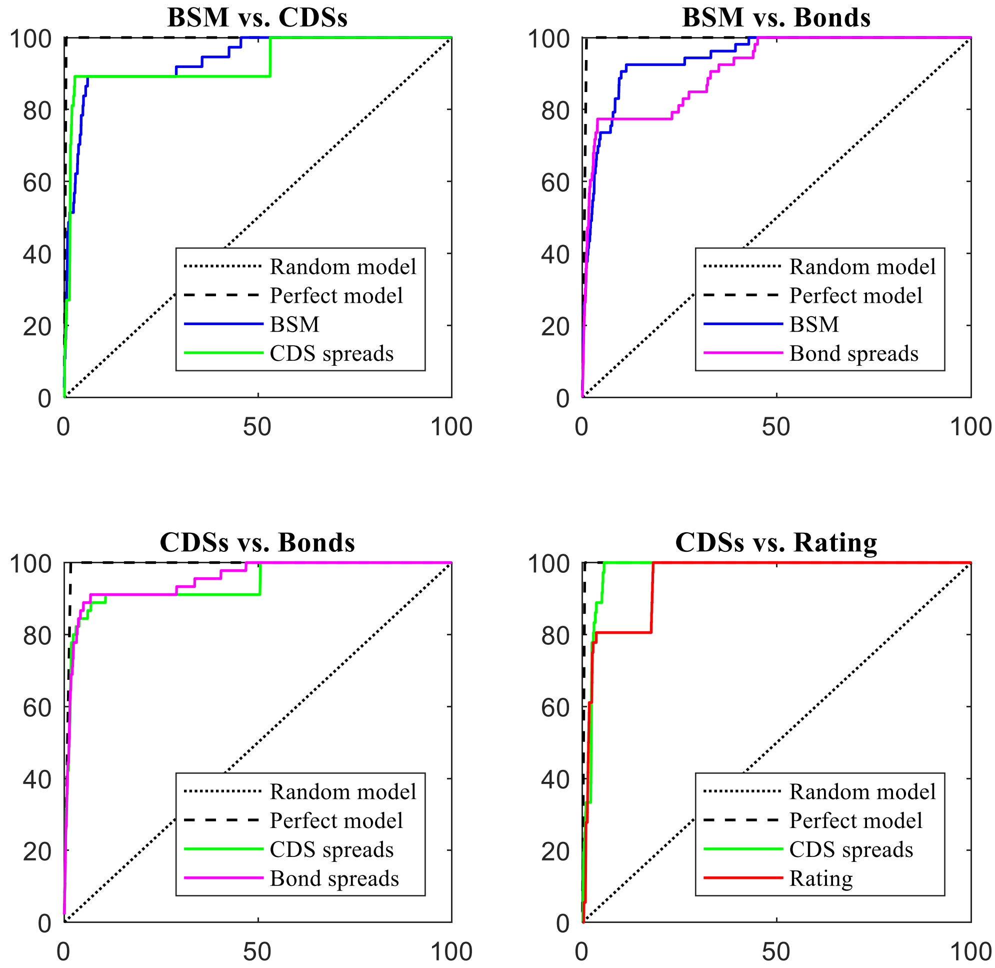


Figure 3: Accumulated matched goodness of fit of credit-risk measures. Default in one year. All default events.

This figure shows the accuracy ratios (AR) of each credit-risk measure when matched to the rest considering default forecasting horizons of one year and taking into account all default events. The number on the right is the aggregate of all the ARs for every measure, obtained by subtracting the negative ARs. As eight measures are considered, there are seven matched pairs, and the maximum aggregated value is 700%.

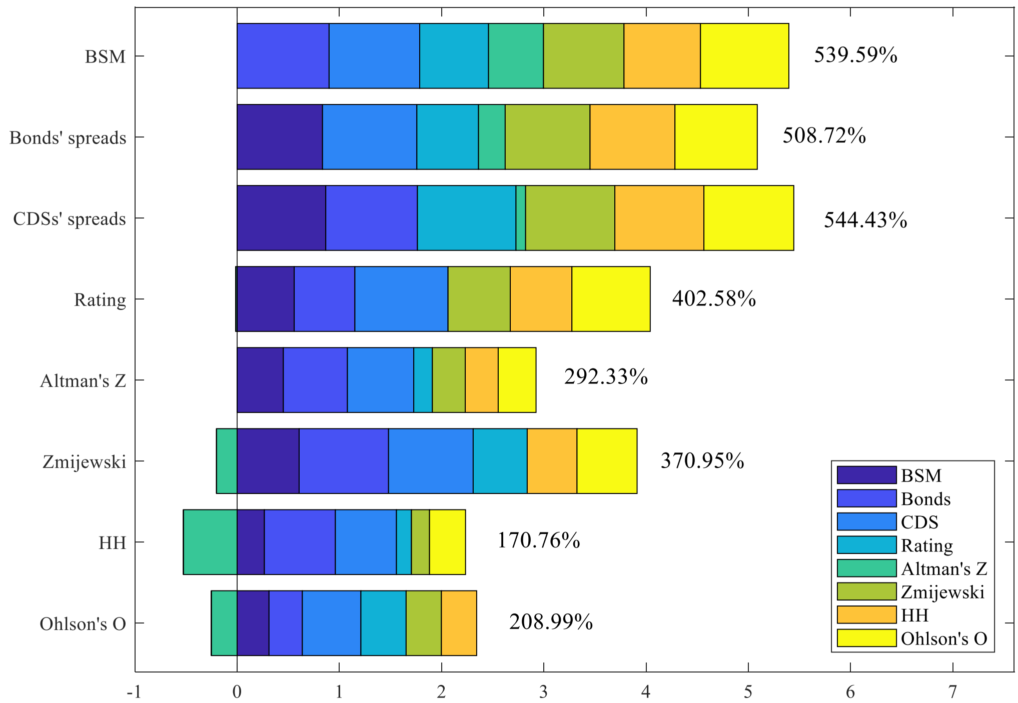


Table 1: Main statistics of credit-risk measures for defaulting and non-defaulting companies.

This table shows the descriptive statistics (mean, maximum, minimum, median and standard deviation) for the credit-risk measures in defaulting and non-defaulting companies. Due to the idiosyncrasy of the credit rating, its median is given instead of the mean. The number of companies with available default data and the number of defaulting companies are shown. The t-test (Z of Wilcoxon's test) shows the difference of means (medians) between defaulted and non-defaulted companies. *** and ** denote significance at the 1 and 5% levels, respectively.

	Number of companies		Defaulted companies					Non-defaulted companies					Tests	
	All	**Defaulted**	**Mean**	**Max.**	**Min.**	**Median**	**SD**	**Mean**	**Max.**	**Min.**	**Median**	**SD**	**t-statistic**	**Z**
BSM	629	84	0.05	1.00	0.00	0.00	0.08	0.02	1.00	0.00	0.00	0.04	3.01***	2.41***
Bond spreads	129	32	444.18	7681.50	10.50	330.88	181.87	358.85	4149.70	10.80	321.05	99.05	1.48	0.88
CDS spreads	138	30	527.05	13366.96	22.26	202.95	414.34	282.59	9673.97	12.58	145.42	150.97	1.73**	2.05**
Rating	459	87	B1	Aaa	WR[23]	B1		Ba2	Aaa	C	Ba2		2.95***	3.10***
Altman's Z	318	49	4.61	405.51	-2.41	2.40	3.48	241.36	434018.47	-3.44	2.92	500.19	-1.07	-1.68**
Zmijewski	509	75	-0.56	186.78	-19.83	-2.21	1.18	-1.90	353.56	-6.92	-2.51	0.92	1.15	2.24**
Hannan and Hanweck	581	80	0.02	1.00	0.00	0.00	0.02	0.01	1.00	0.00	0.00	0.02	0.68	1.81**
Ohlson's O	353	53	-6.80	2.53	-20.46	-6.54	0.77	-6.99	3.16	-47.14	-6.77	0.59	0.61	1.21

[23] Withdrawn rating.

Table 2: Accuracy Ratios for the different credit-risk measures. Non-matched samples.

This table shows the accuracy ratios (AR) of all the credit-risk measures for the non-matched samples, considering default forecasting horizons from one to five years and taking into account all default events (Panel A) and only severe default events (Panel B). AR values greater than 70% are marked in bold and considered a good fit.

Panel A: All default events

Default in	BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
1 year	60.31%	**70.09%**	**90.88%**	41.11%	32.24%	48.70%	15.18%	34.53%
2 years	55.08%	57.02%	**82.30%**	34.79%	27.92%	51.81%	20.25%	19.38%
3 years	52.78%	46.84%	**79.19%**	28.28%	19.67%	46.05%	17.60%	6.84%
4 years	49.84%	43.85%	**77.61%**	21.25%	23.65%	38.23%	14.17%	3.63%
5 years	45.04%	43.25%	**74.39%**	19.09%	25.98%	32.28%	8.99%	-0.70%

Panel B: Only severe default events

Default in	BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
1 year	**82.23%**	**93.70%**	**90.56%**	51.46%	34.70%	45.15%	20.10%	56.37%
2 years	63.91%	**74.68%**	**81.49%**	47.61%	25.63%	43.12%	20.49%	39.04%
3 years	54.03%	64.04%	**78.27%**	44.06%	15.64%	30.15%	11.17%	25.41%
4 years	46.62%	61.81%	**76.68%**	40.86%	21.49%	20.87%	7.01%	18.44%
5 years	39.09%	60.00%	**73.37%**	38.03%	24.67%	13.25%	0.85%	10.99%

Table 3: Accuracy Ratios for the various credit-risk measures. Matched sample. Default in one year. All default events.

This table shows the accuracy ratios (AR), and their differences, for the subset of companies with concurrent data on two specific measures of credit risk. For example, the entry for [BSM, Bond spreads] in the table represents the ARs for BSM and bond spreads, while the value in brackets corresponds to their difference, calculated as the AR of BSM minus the AR of bond spreads. The analysis is performed considering a one-year default-forecasting horizon and taking into account all default events. AR values greater than 70% are marked in bold and considered a good fit.

	Bond spreads	**CDS spreads**	**Rating**	**Altman's Z**	**Zmijewski**	**H-H**	**Ohlson's O**
BSM	**[6.42%]** **89.92%** **83.50%**	**[1.95%]** **88.62%** **86.67%**	[11.54%] 67.33% 55.79%	[8.56%] 53.74% 45.18%	**[18.15%]** **78.75%** 60.60%	**[48.19%]** **74.81%** 26.62%	**[55.25%]** **86.42%** 31.17%
Bond spreads		**[2.61%]** **92.26%** **89.65%**	[0.96%] 60.31% 59.35%	[-36.59%] 26.07% 62.66%	**[-4.47%]** **83.00%** **87.47%**	**[13.61%]** **83.05%** 69.44%	**[47.94%]** **80.53%** 32.59%
CDS spreads			**[5.26%]** **96.39%** **91.13%**	[-55.47%] 9.41% 64.88%	**[4.50%]** **87.27%** **82.77%**	**[27.38%]** **87.13%** 59.75%	**[30.65%]** **87.91%** 57.26%
Rating				[-19.66%] -1.44% 18.22%	[7.98%] 60.84% 52.86%	[45.61%] 60.30% 14.69%	**[32.36%]** **76.61%** 44.25%
Altman's Z					[52.36%] 32.15% -20.21%	[84.83%] 32.23% -52.60%	[62.29%] 37.01% -25.28%
Zmijewski						[31.05%] 48.70% 17.65%	[24.26%] 58.76% 34.50%
H-H							[0.71%] 35.21% 34.50%

Table 4: Median ARs for different horizons and types of default events. Matched samples.

This table shows the median values for the accuracy ratios (AR) obtained for each credit-risk measure when matched to the rest of the measures, considering default forecasting horizons from one to five years and taking into account all default events (Panel A) and only severe default events (Panel B). AR values greater than 70% are marked in bold and considered a good fit.

Panel A: All default events

Default in	BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
1 year	**78,75%**	**83,00%**	**87,27%**	60,30%	37,01%	58,76%	26,62%	34,50%
2 years	65,30%	66,28%	**72,33%**	56,79%	28,16%	59,28%	31,14%	19,29%
3 years	55,18%	58,10%	69,25%	54,95%	19,13%	50,37%	24,00%	6,93%
4 years	46,65%	58,58%	69,34%	52,80%	23,58%	38,58%	19,79%	3,43%
5 years	39,69%	56,48%	66,36%	52,28%	29,06%	32,28%	14,41%	-0,90%

Panel B: Only severe default events

Default in	BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
1 year	**87,95%**	**92,13%**	**87,27%**	**93,77%**	40,53%	52,02%	38,54%	56,33%
2 years	66,29%	69,69%	**72,33%**	**85,94%**	30,00%	48,36%	34,76%	38,93%
3 years	54,35%	62,84%	69,25%	**85,00%**	19,79%	33,92%	18,26%	25,25%
4 years	43,70%	64,06%	69,34%	**85,01%**	23,58%	24,50%	9,63%	18,26%
5 years	34,70%	59,80%	66,36%	**85,25%**	29,06%	17,59%	3,88%	10,79%

Table 5: Average values of main firm characteristics per measure

This table shows the average values for the main firm characteristics of the companies with data for each credit-risk measure. *Size* is the logarithm of market value of equity, *BTM* is the book-to-market ratio, *Volatility* is the equity volatility of the past 12 months and *Intangibility* is the ratio of total intangible assets to total assets.

	Size	**BTM**	**Volatility**	**Intangibility**
BSM	21.26	0.5215	0.3817	0.1924
Bond spreads	22.26	0.5196	0.3288	0.1662
CDS spreads	22.33	0.5000	0.3649	0.1885
Rating	21.47	0.5578	0.3803	0.2089
Altman's Z	21.55	0.5260	0.3822	0.2676
Zmijewski	21.44	0.5150	0.3920	0.2386
H-H	21.45	0.5300	0.3842	0.2135
Ohlson's O	21.56	0.5080	0.3795	0.2365

Table 6: Accuracy ratios by quartiles and halves of firm characteristics. Default in one year. All default events.

Panel A shows the ARs for the first and fourth size, book-to-market ratio, volatility and intangibility quartiles, for the eight credit-risk measures considered. Panel B shows the ARs for the companies under and above the 50th percentile for size, book-to-market ratio, volatility and intangibility, for the eight credit-risk measures considered. AR values greater than 70% are marked in bold and considered a good fit. NA indicates the non-availability of data due to the lack of default events for the corresponding quartiles.

Panel A: By quartiles

		BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
Size	**Q4**	69.77%	1.99%	29.93%	50.22%	67.94%	31.61%	-43.79%	-33.71%
	Q1	**75.82%**	**82.87%**	**72.91%**	42.63%	-8.82%	40.73%	20.67%	43.69%
BTM	**Q4**	**70.97%**	60.03%	57.72%	40.14%	-7.60%	21.57%	-13.95%	-79.11%
	Q1	**72.21%**	**97.66%**	**98.30%**	57.09%	32.89%	60.99%	39.76%	60.56%
Volatility	**Q4**	**72.29%**	**83.02%**	**89.36%**	48.32%	3.18%	53.23%	24.55%	43.86%
	Q1	52.45%	NA	NA	-23.82%	NA	**82.93%**	-74.64%	**90.09%**
Intangibility	**Q4**	**76.80%**	28.50%	**98.17%**	40.48%	18.69%	-18.16%	-44.55%	36.06%
	Q1	**73.16%**	**84.81%**	**79.64%**	**71.87%**	63.47%	**77.07%**	57.12%	**70.78%**

Panel B: By halves

		BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
Size	**H2**	48.33%	-12.61%	9.22%	38.67%	50.29%	20.93%	-38.18%	-64.97%
	H1	66.10%	**76.76%**	**86.63%**	32.28%	0.45%	44.90%	18.58%	33.00%
BTM	**H2**	63.03%	45.08%	67.66%	41.72%	15.64%	27.68%	-15.54%	-66.84%
	H1	61.58%	**75.79%**	**98.76%**	47.85%	40.63%	63.60%	37.42%	65.15%
Volatility	**H2**	68.69%	69.41%	**82.33%**	47.06%	18.68%	47.45%	21.50%	29.71%
	H1	47.17%	-1.34%	**98.55%**	7.67%	58.80%	54.71%	-70.10%	**89.23%**
Intangibility	**H2**	42.06%	31.50%	**98.17%**	31.80%	30.88%	28.77%	-10.90%	39.60%
	H1	**70.63%**	**80.95%**	**84.09%**	62.27%	63.88%	**70.85%**	55.50%	34.06%

Table 7: Average values of the AR and linear dependence of AR on firm characteristics in the bootstrapped sample. Default in one year. All default events.

Panel A shows the mean and standard deviation of the AR obtained through the bootstrap procedure for each credit-risk measure. Mean AR values greater than 70% are marked in bold and considered a good fit. Panel B shows the results of the OLS estimation. The dependent variable is the Accuracy Ratio (AR) for each measure of credit risk. *Size* is the logarithm of the market value of equity, *BTM* is the book- to-market ratio, *Volatility* is the equity volatility of the past 12 months and *Intangibility* is the ratio of total intangible assets to total assets. *** and ** denote significance at the 1 and 5% levels, respectively.

Panel A: Mean and standard deviation of AR values for bootstrapped samples

	BSM	Bond spreads	CDS spreads	Rating	Altman's Z	Zmijewski	H-H	Ohlson's O
Mean	60.60%	65.83%	**87.33%**	42.80%	40.47%	39.13%	-1.91%	32.73%
Std.	27.71%	26.61%	16.45%	27.72%	25.13%	36.67%	52.08%	38.44%

Panel B: Linear dependence of AR on firm characteristics in bootstrapped samples

	Constant	Size	BTM	Volatility	Intangibility
			Panel B.1: Model 1		
BSM	3.0224***	-0.1101**	-0.2144**		
Bond spreads	13.6061***	-0.5849***	0.5791***		
CDS spreads	4.8472***	-0.1685***	-0.4463***		
Rating	-1.6953**	0.1016***	-0.0471		
			Panel B.2: Model 2		
BSM	0.50214	-0.0174	-0.0046	1.1461***	
Bond spreads	9.8462***	-0.4468***	0.4453***	1.9273***	
CDS spreads	4.2963***	-0.1531***	-0.3463***	0.4489**	
Rating	-1.9619*	0.1116***	-0.0356	0.1271	
			Panel B.3: Model 3		
BSM	3.0757***	-0.1089**	-0.2127**		-0.4415
Bond spreads	9.7149***	-0.3824***	0.1634		-2.4955***
CDS spreads	4.6662***	-0.1622***	-0.5137***		0.3653
Rating	-1.6376**	0.1004***	-0.0521		-0.1685

Table 8: Average values of the firm characteristics in the bootstrapped sample. Default in one year. All default events.

This table shows the average values of the main firm characteristics of those companies in the bootstrapped samples with data available for each measure of credit risk. *Size* is the logarithm of the market value of equity, *BTM* is the book-to-market ratio, *Volatility* is the equity volatility of the past 12 months and *Intangibility* is the ratio of total intangible assets to total assets.

		Size	BTM	Volatility	Intangibility
BSM	**Mean**	20.8987	0.5345	0.4106	0.1779
	Std.	0.2606	0.112	0.0308	0.0226
Bond spreads	**Mean**	22.6287	0.4993	0.3637	0.1939
	Std.	0.2548	0.0484	0.0277	0.0392
CDS spreads	**Mean**	22.2697	0.4952	0.3531	0.2015
	Std.	0.1403	0.0617	0.0264	0.0286
Rating	**Mean**	21.1654	0.5919	0.3856	0.174
	Std.	0.2885	0.0842	0.0329	0.0258
Altman's Z	**Mean**	21.8231	0.5131	0.372	0.2706
	Std.	0.1705	0.0562	0.0214	0.0278
Zmijewski	**Mean**	21.4396	0.5108	0.3906	0.2281
	Std.	0.2668	0.049	0.0197	0.0249
H-H	**Mean**	21.4569	0.5232	0.3834	0.2077
	Std.	0.2666	0.0471	0.0205	0.0249
Ohlson's O	**Mean**	21.6810	0.4819	0.3712	0.2343
	Std.	0.1427	0.0481	0.0148	0.0235

Table 9: Linear dependence of AR on firm characteristics in the bootstrapped sample restricted to companies with CDS data. Default in one year. All default events.

This table shows the results of the OLS estimation. The dependent variable is the Accuracy Ratio (AR) for each measure of credit risk when matched with CDS spreads. *Size* is the logarithm of the market value of equity, *BTM* is the book-to-market ratio, *Volatility* is the equity volatility of the past 12 months and *Intangibility* is the ratio of total intangible assets to total assets. *** and ** denote significance at the 1 and 5% levels, respectively.

	Constant	Size	BTM	Volatility	Intangibility
Panel A: Model 1					
BSM	8.3420***	-0.3126***	-1.0796***		
Bond spreads	4.7354***	-0.1612***	-0.4489***		
CDS spreads	4.8472***	-0.1685***	-0.4463***		
Rating	-0.0508	0.0551***	-0.5677***		
Panel B: Model 2					
BSM	8.4701***	-0.3096***	-1.1974***	-0.3921	
Bond spreads	3.4783***	-0.1146***	-0.4405***	0.5344***	
CDS spreads	4.2963***	-0.1531***	-0.3463***	0.4489**	
Rating	-0.6888	0.0779***	-0.5768***	0.3542	
Panel C: Model 3					
BSM	7.8112***	-0.2909***	-1.2490***		0.6649***
Bond spreads	4.0741***	-0.1363***	-0.5370***		0.7959***
CDS spreads	4.6662***	-0.1622***	-0.5137***		0.3653
Rating	0.2467	0.0398	-0.5785***		0.2791**

Table 10: Accuracy ratios by quartiles of firm characteristics and linear dependence of AR on firm characteristics in the bootstrapped sample. Default in one year. All default events.

Panel A shows the ARs for the first and fourth quartiles for the newly-added firm characteristics, that is, profitability, liquidity and operating cycle. AR values greater than 70% are marked in bold and considered a good fit. NA indicates the non-availability of data due to the lack of default events for the corresponding quartiles. Panel B shows the results of the OLS estimation. The dependent variable is the Accuracy Ratio (AR) for each measure of credit risk. *Size* is the logarithm of the market value of equity, *BTM* is the book-to-market ratio, *Volatility* is the equity volatility of the past 12 months, *Liquidity* is cash divided by total assets, *Profitability* is net income divided by total assets and *Operating Cycle* is the average period of time between the outlay of cash to produce a product and the receipt of cash from the sale of the product. *** and ** denote significance at the 1 and 5% levels, respectively.

Panel A: Accuracy ratios by firm characteristic quartiles

		BSM	Bond spreads	CDS spreads	Rating
Liquidity	**Q4**	**86.76%**	**80.35%**	**72.86%**	**82.57%**
	Q1	68.32%	**85.30%**	**96.96%**	**73.28%**
Profitability	**Q4**	**93.43%**	68.28%	NA	45.74%
	Q1	**73.24%**	**93.91%**	**89.01%**	59.22%
Operating cycle	**Q4**	**86.42%**	68.28%	**94.67%**	45.74%
	Q1	**74.51%**	**93.91%**	**88.86%**	59.22%

Panel B: Linear dependence of AR on firm characteristics in the bootstrapped samples

	Constant	Size	BTM	Liquidity	Profitability	Operating cycle
Panel B.1: Model 4						
BSM	3.4016***	-0.1381***	-0.1978**	2.8953***		
Bond spreads	13.7516***	-0.5901***	0.5731***	-0.5239		
CDS spreads	4.6635***	-0.1661***	-0.4380***	2.1893***		
Rating	-1.6340**	0.1023***	-0.0604	-1.3447		
Panel B.2: Model 5						
BSM	3.0058***	-0.1089**	-0.2134**		-0.2430	
Bond spreads	13.5978***	-0.5844***	0.5771***		-0.0808	
CDS spreads	4.0568***	-0.1294***	-0.4193***		-2.4196**	
Rating	-1.6846**	0.1039***	-0.0551		-1.3808	
Panel B.3: Model 6						
BSM	3.1277***	-0.1129***	-0.2151**			-0.0004
Bond spreads	13.9629***	-0.5959***	0.5443***			-0.0009
CDS spreads	4.4411***	-0.1539***	-0.4843***			0.0010
Rating	-1.7180**	0.1022***	-0.0483			0.0001

Table 11: Summary of results for the predictive accuracy of credit-risk measures as a function of firm characteristics.

This table shows a summary of the main conclusions to be drawn from the analysis of the results relating to the predictive accuracy of the credit-risk measures as a function of sample characteristics. The analysis is performed distinguishing cases between a default-forecasting horizon of one year and of five years, and between performing the analysis for all default events and considering severe default events.

		All default events			Severe default events		
		Overall predictive power	Predictive power according to characteristics		Overall predictive power	Predictive power according to characteristics	
			Effect	No effect		Effect	No effect
Default in one year	**BSM**	Good fit	High predictive power for small firms and firms with high volatility	BTM, intangibility, liquidity, profitability, operating cycle	Good fit	High predictive power for small firms and firms with low BTM	Volatility, intangibility, liquidity, profitability, operating cycle
	Bonds spreads	Good fit	High predictive power for small firms, firms with low BTM, high volatility and low intangibility	Liquidity, profitability, operating cycle	Good fit	High predictive power for small firms and firms with low BTM	Size, volatility, liquidity, profitability, operating cycle
	CDS spreads	Good fit	High predictive power for small firms and firms with low BTM	Volatility, intangibility, liquidity, profitability, operating cycle	Good fit	High predictive power for small firms and firms with low BTM	Volatility, intangibility, liquidity, profitability, operating cycle
	Rating	Poor fit, except for liquidity	No effect	Size, BTM, volatility, intangibility, liquidity, profitability, operating cycle	Good fit	No effect	Size, BTM, volatility, intangibility, liquidity, profitability, operating cycle
	Accounting measures	Poor fit	Not analysed due to the poor goodness of fit of the measures	Not analysed due to the poor goodness of fit of the measures	Poor fit	Not analysed due to the poor goodness of fit of the measures	Not analysed due to the poor goodness of fit of the measures
Default in five years	**BSM**	Poor fit, except for liquidity	No effect	Size, BTM, volatility, intangibility, liquidity, profitability, operating cycle	Good fit for BTM, intangibility, liquidity and operating cycle	High predictive power for firms with high operating cycle	Size, BTM, volatility, intangibility, liquidity, profitability
	Bonds spreads	Good fit for intangibility, liquidity, profitability and operating cycle	High predictive power for firms with low intangibility, low liquidity and low profitability	Size, BTM, volatility, operating cycle	Good fit for BTM, intangibility, liquidity, profitability and operating cycle	High predictive power for firms with low BTM, low intangibility, low profitability and high operating cycle	Size, volatility, liquidity
	CDS spreads	Good fit	High predictive power for small firms, firms with low BTM, high volatility and low profitability	Intangibility, liquidity, operating cycle	Good fit	High predictive power for small firms and firms with low BTM and low profitability	Volatility, intangibility, liquidity, operating cycle
	Rating	Poor fit	No effect	Size, BTM, volatility, intangibility, liquidity, profitability, operating cycle	Good fit for intangibility, liquidity and profitability	No effect	Size, BTM, volatility, intangibility, liquidity, profitability, operating cycle
	Accounting measures	Poor fit	Not analysed due to the poor goodness of fit of the measures	Not analysed due to the poor goodness of fit of the measures	Poor fit	Not analysed due to the poor goodness of fit of the measures	Not analysed due to the poor goodness of fit of the measures

Table 12: Results of the regression of the accuracy ratio over the characteristics of the sample for all eight measures at once. Default in one year. All default events.

This table shows the results of the OLS estimation for all the credit-risk measures simultaneously. The dependent variable is the Accuracy Ratio (AR), *Size* is the logarithm of market value of equity, *BTM* is the book to market ratio, *Volatility* is the equity volatility of the past 12 months, *Liquidity* is cash divided by total assets, *Profitability* is the net income divided by total assets and *Operating Cycle* is the average period of time between the outlay of cash to produce a product and the receipt of cash from the sale of the product. *D_h* is the dummy variable that takes value one when the credit-risk measure is *h* and 0 otherwise. *** and ** denote significance at the 1 and 5% levels, respectively.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	2.8211***	0.9228	2.3892***	2.8066***	2.7981***	2.8587***
Size	-0.0828***	-0.0169	-0.0529***	-0.0858***	-0.0815***	-0.0898***
BTM	-0.2075***	-0.1052	-0.1862***	-0.1961***	-0.2080***	-0.2361***
Volatility		1.0724***				
Intangibility			-1.2188***			
Liquidity				1.3067***		
Profitability					-0.1632	
Operating cycle						0.0013***
D_BSM	-0.3726***	-0.3480***	-0.3611***	-0.3903***	-0.3701***	-0.3888***
D_Bond	-0.1843***	-0.2197***	-0.2042***	-0.1724***	-0.1836***	-0.1818***
D_Rating	-0.5166***	-0.4886***	-0.5191***	-0.5117***	-0.5151***	-0.5268***
D_AltmanZ	-0.5018***	-0.4939***	-0.4033***	-0.5295***	-0.5000***	-0.4963***
D_ZM	-0.5473***	-0.5346***	-0.4903***	-0.5733***	-0.5455***	-0.5448***
D_HH	-0.9538***	-0.9355***	-0.9224***	-0.9738***	-0.9524***	-0.9521***
D_OhlsonO	-0.5975***	-0.5769***	-0.5397***	-0.6247***	-0.5949***	-0.5966***