# On the prediction of financial distress for UK firms: Does the choice of accounting and market information matter?

# On The Prediction Of Financial Distress For UK firms: Does the Choice of Accounting and Market Information Matter?

Evangelos C. Charalambakis*
Susanne K. Espenlaub†
Ian Garrett‡

*Corresponding author. University of Central Greece, Lambrou Katsoni Square, Livadeia 32 100, Greece.
†Manchester Business School, University of Manchester, UK, email: susanne.espenlaub@mbs.ac.uk
‡Manchester Business School, University of Manchester, email: ian.garett@mbs.ac.uk. We would like to thank Vasiliki Athanasakou, Asad Kausar and the participants at the 2009 EFMA Conference in Milan for helpful comments.

# On The Prediction Of Financial Distress For UK firms: Does the Choice of Accounting and Market Information Matter?

**Abstract**

We assess the contribution of accounting and market-driven variables to the prediction of bankruptcy for UK firms. Using the hazard approach recommended by Shumway (2001), we show that a hazard model that combines both accounting *and* market information provides more accurate predictions of the probability of financial distress than the accounting ratio-based Z-score model and models exclusively based on market-related covariates. When we decompose the Z-score into its component ratios we find that in the hazard model, half of them do not contribute to the prediction of corporate failure. When we incorporate Z-score as an additional covariate in a model that also contains market information, the Z-score does not possess any incremental predictive power. Finally, a comparison of the ability of different information sets (accounting information only, market information only and a combination of the two) to predict financial distress both in-sample and out-of-sample shows that a hazard model combining both accounting and market-driven variables generates the best performance in terms of predictive ability, while a model based exclusively on market-driven variables outperforms that based solely on accounting variables. Our results suggest that for the UK at least, there are significant gains to be had from predicting financial distress using both accounting and market information.

# 1 Introduction

Academics and practitioners have long shown interest in the prediction of corporate bankruptcy and financial distress.[1] Since the seminal work of Beaver (1966) and Altman (1968) researchers have, until recently, almost without exception developed bankruptcy prediction models using accounting information, particularly accounting ratios, as the variables or covariates that predict financial distress.[2] More recent studies, particularly for the US, have focused on how models based on market information, in particular the Merton (1974) structural model for pricing corporate debt, perform in predicting bankruptcy (Duffie, Saita and Wang (2007) and Bharath and Shumway (2008) for the US, and Agarwal and Taffler (2008) for the UK, for example) while Shumway (2001) and Campbell, Hilscher and Szilagyi (2008) consider the interaction of both accounting and market information for the US. Evidence from Shumway (2001) and Campbell, Hilscher and Szilagyi (2008) shows that, for the US at least, a combination of both accounting and market information provides more accurate predictions of the probability of bankruptcy. For the UK there is a paucity of studies examining this issue.[3] We remedy this in this paper. We explore the extent to which accounting and/or market-driven information is associated with the prediction of financial distress. Using the model and approach suggested by Shumway (2001) we assess the performance of a bankruptcy prediction model for UK firms using accounting data, market data and a combination of the two. This is an important line of enquiry as it will shed light on whether market information acts as a substitute or complement to accounting information in predicting bankruptcy. This in turn is important because measures of financial distress are increasingly used in asset pricing papers that address the issue of whether distress risk is priced (Dichev (1998), Griffin and Lemmon (2002) and Campbell, Hilscher and Szilagyi (2008)) and empirical research on capital structure, where the Z-score is typically used (see, for example, Graham (2000) and Byoun (2008) among others.) Clearly evaluating the accuracy of predictions of financial distress and

[1] For the remainder of the paper, unless otherwise indicated, we use the terms bankruptcy and financial distress interchangeably.

[2] See, among many others, Ohlson (1980), Zmijewski (1984) and Beaver, McNichols and Rhie (2005) for the US and Taffler (1983) and Agarwal and Taffler (2007) for the UK.

[3] Agarwal and Taffler (2008) compare Taffler's Z-score model (Taffler (1983)) with the KMV-Merton-based models used by Hillegeist, Keating, Cram and Lundstedt (2004) and Bharath and Shumway (2008) using UK data but they do not examine whether a model combining features of both provides more accurate predications of financial distress.

the role of accounting and market-based information in enhancing such forecasts is of great importance and interest.

There are good reasons to argue that accounting ratio-based models are less informative than market-based models with respect to bankruptcy prediction. First, accounting-based models use information from financial statements, which verify the firm's past performance while providing only limited cues about the firm's future status. Second, accounting conventions (historical cost and conservatism, for example) limit the scope of accounting information so that the book value of assets in the financial statements is usually understated. Third, while accounting data provide a snapshot of the value of the company at a specific point in time, market data is forward-looking. However, the forward-looking nature of market information does not necessarily suggest that market-related variables can subsume accounting ratios when predicting corporate financial distress.

In addition to the information set another important element for the prediction of bankruptcy is the methodology used to estimate its likelihood. Prior literature uses various methods. Altman (1968) uses discriminant analysis to estimate Z-scores, while Ohlson (1980) and Zmijewski (1984)use logit and probit models to predict the probability of financial distress. There are, however, methodological issues that arise in using these models. Shumway (2001) argues that well-established bankruptcy prediction models, such as Altman's (1968) Z-score and Ohlson's (1980) conditional logit model are not correctly specified as they do not consider all the available firm-year observations. This induces a bias in the estimated coefficients on the variables used to predict bankruptcy, leading to incorrect statistical inferences. Shumway (2001) develops a hazard model that takes into account all the available observations for the bankrupt and non-bankrupt firms and that overcomes problems relating to biased parameter estimates and statistical inference. He shows that using the hazard model delivers efficient and consistent parameter estimates. In addition, he documents that when using a hazard model, half of the accounting ratios incorporated in Altman's (1968) and Zmijewski's (1984) accounting-based models are not statistically significant for predicting bankruptcy. He proposes a hazard model that uses both accounting ratios and market-driven variables as this outperforms the two accounting-based models using Altman's and Zmijewski's variables in out-of-sample forecasts. This suggests that there is some mileage in using a combination of accounting information, giving a snapshot of the firm's current financial health, and market information, providing a view of where the market thinks the firm may be heading, to forecast the probability of bankruptcy.

We begin our investigation using a reduced-form hazard model in the spirit of Shumway (2001).[4] We divide the information set from which we draw our predictor variables into three categories. First, we examine the ability of accounting ratios to predict the probability of financial distress. We begin by examining whether the popular Z-score measure, which is calculated using accounting ratios, contains any predictive information about the probability of bankruptcy. Given that we are interested in predicting the probability of financial distress for UK firms, we calculate the Z-score using Taffler's model (Taffler (1983) and Agarwal and Taffler (2007)). Given that the Z-score and the probability of financial distress should be capturing the same thing, we would expect to find a significant relationship between them. We also "decompose" the Z-score into it's individual components to examine which of the accounting ratios that comprise the Z-score significantly predict the probability of bankruptcy. Second, we examine the ability of the market-based variables used by Shumway (2001) to predict bankruptcy for UK firms. Third, we investigate the role of both accounting *and* market-based information in predicting bankruptcy. Encouragingly, we find that in a univariate model, Z-score does significantly predict the probability of financial distress. However, when the Z-score is decomposed into its individual components we find that only two of the four accounting ratios that make up the Z-score are statistically significantly different from zero. Things deteriorate rapidly for accounting ratios based on the Z-score when faced with market information, however. We find that excess returns, relative size, based on market capitalization, and return volatility significantly predict the probability of bankruptcy. Further, when we include the Z-score in the model with market-based variables and the individual components comprising the Z-score in the model with market-based variables, they have no incremental predictive power above that contained in the market-based variables. Given the apparent inability of the Z-score and its component ratios to predict bankruptcy once excess returns, volatility and relative size are considered, we broaden the set of accounting ratios and find that in addition to the market-based variables, book leverage is statistically significant.

[4]We use the term reduced form here to distinguish a model using predictors found to be significant in other studies from those derived from Merton's (1974) structural model.

Our results therefore suggest that a combination of accounting and market-based variables best (in terms of statistical significance) describe the probability of bankruptcy. Our findings are confirmed by a series of in-sample and out-of-sample prediction tests. These show that both the hazard model using accounting and market-based variables, and the hazard model using market-based variables alone dominate one based on the accounting ratios that comprise Taffler's Z-score. Also, consistent with Shumway (2001) we find that the accounting and market-based model outperforms the market-based model alone.

Our results have important implications. In particular, they suggest that researchers and analysts should use a hazard model that combines both accounting and market-based information rather than a model that relies exclusively on accounting ratios or market-related variables. The amalgam of accounting and market information yields more accurate predictions of corporate bankruptcy. The rest of the paper is organized as follows. Section 2 provides an overview of how we model the probability of financial distress. Section 3 describes our data. Section 4 presents the results from the various bankruptcy forecast models. Section 5 reports the in-sample and out-of-sample forecast accuracy of the models while section 6 concludes.

## 2 Modeling the probability of financial distress

A variety of estimation techniques have been employed to develop bankruptcy forecasting models. Beaver (1966) uses a multiple regression model to predict corporate failure with accounting ratios. Altman (1968) employs multivariate discriminant analysis to derive the Z-score measure for predicting bankruptcy for US firms; Taffler (1983) uses the same technique for UK firms, while Altman, Haldeman and Narayanan (1977) use quadratic discriminant analysis to identify firms in danger of going bankrupt. Ohlson (1980) estimates a conditional logit model to generate the probability that a firm will enter bankruptcy (known as the "O-score") while Zmijewski (1984) estimates a probit model. Lau (1987) uses a multinomial logit model that allows for more than two states of financial distress. Shumway (2001), however, argues that these bankruptcy forecasting models are misspecified as they do not properly account for the length of time that a healthy firm has survived. In particular, such models only use observations on the explanatory variables for a single firm-year (hence they can be thought of as static), which can be arbitrarily chosen for non-bankrupt firms, while for non-bankrupt firms the firm-year observation is not randomly selected, typically corresponding to the year before bankruptcy. This induces a selection bias (Shumway (2001), Hillegeist, Keating, Cram and Lundstedt (2004)). Shumway (2001) shows that ignoring information about the length of time a healthy firm has survived produces biased and inconsistent estimates of the parameters of the model. To properly address this, Shumway (2001) uses a discrete time hazard model. In the hazard model, the hazard rate is the probability of the firm going bankrupt at time $t$ conditional upon having survived until time $t$. Therefore, in a hazard model the probability of bankruptcy changes through time. This variation in the probability of bankruptcy over time not only allows researchers to take advantage of all the available firm-year observations, it allows the probability of bankruptcy to change as a function of a vector of explanatory variables, known as covariates, that also change over time. The general form of the hazard model is

$$\ln\left[\frac{h_i(t)}{1-h_i(t)}\right] = \alpha(t) + \boldsymbol{\beta}'\mathbf{x}_{it} \tag{1}$$

where $h_i(t)$ represents the hazard of bankruptcy at time $t$ for firm $i$, conditional on survival to $t$; $\alpha(t)$ is the baseline hazard; $\boldsymbol{\beta}$ is a vector of coefficients and $\mathbf{x}_{i,t}$ a $k \times 1$ vector of observations on the $i$th covariate at time $t$. The attraction of this approach, as Shumway (2001) shows, is that the discrete-time hazard model is econometrically equivalent to a dynamic logit model where each period that a firm survives is included as a non-failing firm. Therefore, we estimate the probability of bankruptcy as

$$P_{t-1}(Y_{it} = 1) = \frac{1}{1 + \exp\left(-\alpha - \boldsymbol{\beta}'\mathbf{x}_{it-1}\right)} \tag{2}$$

where $Y_{it}$ is a variable that equals one if firm $i$ goes bankrupt in year $t$, zero otherwise. $\boldsymbol{\beta}$ and $\mathbf{x}$ are as before. Notice that we use data dated $t-1$ in estimating the probability of bankruptcy. This is to ensure that we only use data that is actually available at the beginning of the year in which bankruptcy occurs. Given that it is possible to set the hazard model up as a logit model, it is very easy to estimate with one proviso. Before any inference can be undertaken in relation to the significance or otherwise of the elements of $\boldsymbol{\beta}$, it is necessary to adjust the Wald statistic that tests the significance of the coefficients. The reason for this is that because we treat each firm-year observation as if it were a separate firm, estimation using standard logit routines treats the model as if it were static. In the static logit model, the number of firm-years is used in calculating the Wald statistics. However, this is not correct for the dynamic logit model because in the dynamic logit model, unlike the static logit model, firm-year observations are not independent of each other. For the dynamic logit model, it is the number of firms rather than the number of firm years that should be used. The test statistics therefore need to be scaled by the average number of firm years per firm.[5]

## 3 Data

We obtain the accounting data from Datastream and the market data from the London Share Price Database (LSPD). The sample consists of 3,459 (alive and dead) UK listed firms over the period 1980–2006. We exclude financial firms and utilities from the sample. We identify bankrupt firms as follows. For each year in the sample, we search the London Share Price Database (LSPD) database based on the LSPD death type. We define a firm as bankrupt if a firm's LSPD death type is liquidation, voluntary liquidation, receiver appointed/liquidation, in administration/administrative receivership, and canceled assumed valueless, otherwise it is classed as non-bankrupt. We identify 310 bankrupt firms who provide 2,378 firm-year observations in total. There are 3,149 non-bankrupt firms in our sample, providing 29,879 firm-year observations. This gives a total of 32,257 firm-year observations initially, although the actual number of observations available to us to estimate the various dynamic logit models differs according to the data availability relating to each variable; we will return to this point below. Table 1 provides detailed information on the definition of all variables used in the study. Of the accounting variables we use, profit before tax (PROF), working capital (WCAP), financial risk (FRISK) and Liquidity (LIQUID) are the accounting ratios on which Taffler (1983) bases his Z-score. We also include an alternative measure of performance based on earnings (EBITDA_TA) and book leverage (BLEV) in the accounting-variable information

**INSERT TABLE 1 ABOUT HERE**

[5] Define $\boldsymbol{\theta} = [\alpha\ \boldsymbol{\beta}\ ]'$. Define the hypothesis we wish to test as $H(\boldsymbol{\theta})$. In our case, we are interested in whether this is equal to zero. The Wald statistic is $W = nH(\widehat{\boldsymbol{\theta}})'\widehat{\boldsymbol{\Sigma}}^{-1}H(\widehat{\boldsymbol{\theta}})$, where hats denote estimates and $\boldsymbol{\Sigma}$ is the parameter covariance matrix. This test is distributed as $\chi^2(r)$, where $r$ is the number of restrictions. For the static logit model, $n$ is the number of firm-years. For the dynamic logit model, $n$ should be the number of firms. Therefore, scaling the test statistics reported in the logistic regression output by the average number of firm-years per firm will deliver the correct test statistic.

set. The market-driven variables we use are relative size (REL_SIZE), which expresses the equity market capitalization of the firm relative to total equity market capitalization, excess stock returns (EXRET) and idiosyncratic stock return volatility ($\sigma$). Both Shumway (2001) and Campbell, Hilscher and Szilagyi (2008) find that these variables are statistically significant predictors of the probability of financial distress for US firms. We include them here to examine whether these variables are robust predictors of the probability of financial distress. If they are, we would expect their UK equivalents to be significant predictors of the probability of bankruptcy for UK firms. With the exception of REL_SIZE, which appears on the basis of normality tests to be normally distributed, and Z-score, we truncate the independent variables at the 1st and 99th percentiles of the distribution to avoid outliers. For the Z-score, we follow Agarwal and Taffler (2008) and winsorize Z-score so that it lies between $\pm$ 18.4207. Descriptive statistics for the explanatory variables are provided in Table 2.

**INSERT TABLE 2 ABOUT HERE.**

One point to note about the data, and this can be seen in Table 2, is that the number of firm-year observations differs across the variables. This is because some of the accounting and/or the market data is not always available for all firms in all years. In terms of estimating our dynamic logit models, this means that the number of observations available differs not only across accounting-based and market-based information sets but even across information sets that combine accounting and market-based information since the firms that have missing accounting data and missing market data need not be the same: some firms that have accounting data do not have market data and vice-versa; for some firms, not all of the accounting data needed to calculate the ratios is available. We will return to this issue in the next section. We now turn our attention to the results.

# 4 Results

**INSERT TABLE 3 ABOUT HERE.**

The results from estimating a series of dynamic logit models using different information sets are presented in Panel A of Table 3. The results in Table 3 come from models that use all of the available observations for the particular information set. To ensure that our results are not driven by differences in the sample composition, we repeat our analysis using a balanced sample containing the same firm-year observations for each model. The results are qualitatively and quantitatively similar, so we only discuss the results using all available observations here. Results for the balanced sample can be found in the Appendix. We initially focus attention on the results in the first three columns. The column headed ZSCORE is for a model where the only predictor of the probability of financial distress is the Z-score. The column headed ZDECOMP is a multivariate model where we decompose the Z-score into its component accounting ratios to examine which, if any, of the ratios that make up the Z-score individually help predict the probability of financial distress. The column headed MV is a model using the market-driven variables as predictors of financial distress.

Focusing on the ZSCORE column, encouragingly, but perhaps not surprisingly given that the Z-score should be capable of identifying firms in financial distress, we find that the Z-score is a statistically significant predictor of the probability of financial distress and has the correct sign: the higher is the Z-score, the less likely the firm is to fail, and this is what we observe. This is an encouraging start, but if Z-score is a powerful predictor of financial distress, as the results in the ZSCORE column suggest it might be, we might expect all of its component ratios to predict financial distress. The ZDECOMP column reports the results from examining which of the components of the Z-score best predict financial distress. When we break the Z-score down into its constituent parts, we find that the significance of the Z-score in predicting the probability of financial distress is driven by profitability and financial risk, with more profitable firms less likely to fail while firms with higher financial risk (higher current liabilities relative to total assets) more likely to fail. Working capital and liquidity appear to have no incremental predictive power over that contained in profitability and financial risk. The predictive ability of Z-score, then, derives from firms' profitability and financial risk. Turning our attention to the market-based variables, the model in the MV column predicts the probability of financial distress using relative size, excess returns and idiosyncratic return volatility. Size, as measured by relative market capitalization, and excess returns are significant, although size is only significant at the 10% level, and are negatively related to failure while volatility is significantly positively related to failure. The results thus far, then, suggest that on their own, the accounting ratios, or at least some of them, that comprise Taffler's Z-score, and market-based variables are separately capable of predicting the probability of bankruptcy. This raises the intriguing question as to whether one set of variables is better than the other in terms of predicting the probability of financial distress, or whether combining them into one information set will lead to a model that outperforms the ZSCORE, ZDECOMP and MV models. To investigate this question, we follow Hillegeist, Keating, Cram and Lundstedt (2004) and make use of Vuong's (Vuong (1989)) test for model selection between two non-nested models, $i$ and $j$. Under the null hypothesis that there is no difference between the two models, the log of the ratio of the likelihood for model $i$ to that for model $j$ should be zero. If the difference is significantly positive, $i$ is preferred to $j$ and vice versa. Vuong (1989) derives a statistic that allows us to test this hypothesis. Under the null hypothesis that there is no difference between the competing models, the test statistic has a standard normal distribution. Panel B of Table 3 contains results of the Vuong test for various models. Of interest here is how the two Z-score-based models, ZSCORE and ZDECOMP, perform against the model using market-driven predictors, MV. The null hypothesis that there is no difference between MV and ZSCORE, and MV and ZDECOMP is soundly rejected in both cases in favor of the alternative in both cases that MV performs better. As a robustness check on the results from the Vuong test, we supplemented the MV model with the variables from the ZSCORE model and the ZDECOMP model and re-estimated it. Parameter estimates are reported in the columns headed MVZSCORE and MVZDECOMP. In both models, the Z-score-based variables are statistically insignificant: it seems that the predictive power of the accounting ratios used in the Z-score is subsumed by market-related variables.

The finding above that accounting ratios, at least those used in calculating the Z-score, contain no incremental information is counter to evidence from US studies that use market-based and accounting ratios together. Shumway (2001) and Campbell, Hilscher and Szilagyi (2008) find that either or both of some measure of net income to total assets, and a leverage measure, coupled with the market-driven variables, significantly predict the probability of financial distress. To assess whether this is a robust result that applies to a sample other than the US, we estimate a dynamic logit model that includes Earnings before Interest, Tax, Depreciation and Amortization, scaled by Total Assets, and book leverage, along with relative size, excess past returns and idiosyncratic volatility, as predictor variables. Results for this model can be found in the MVACC column of Table 3. While the earnings variable is statistically insignificant, book leverage is significant at the 5% level. The market-driven variables retain their earlier significance. Adding Z-score to this model shows that it has no incremental predictive power, as can be seen in the MVACCZSCORE column.[6] The Vuong tests in Panel B show that the model with market variables and earnings and book leverage is preferred to a model with market-driven variables alone and is also preferred to models incorporating Z-score.

## 5 In-sample and Out-of-sample Forecast Accuracy

INSERT TABLE 4 ABOUT HERE

### 5.1 In-sample Performance

To evaluate the predictive ability of the models from the previous section, we follow Shumway (2001) and sort firms into deciles based on the probability of bankruptcy estimated by the relevant model.[7] Deciles one through five contain firms that are more likely to go bankrupt, decile one containing those firms with the highest predicted probability of bankruptcy, while deciles six through ten contain those firms that are considered least likely to go bankrupt, decile ten containing those firms with the lowest predicted probability of bankruptcy. We then calculate the percentage of bankrupt firms that are allocated to the various deciles.[8] Table 4 presents the results for in-sample predictive ability. This can be thought of as a means by which we can assess the explanatory capability of the various models since we use the entire sample to estimate the models and assess their ability to correctly classify those firms that went bankrupt as likely to go bankrupt. We consider the ability of the models to predict bankruptcy out-of-sample in the next subsection.

An interesting point to note is that all of the models perform well in terms of identifying firms more likely to go bankrupt than less. The ZSCORE model correctly places around 80% of those firms that do go bankrupt in deciles 1 through 5, those firms that are more predicted more likely than not to go bankrupt. Interestingly, although the results in Table 3 show that when we decompose the Z-score into its component ratios only two of the four

[6] We also supplemented the MVACC information set with the individual components of the Z-score. However, given the nature of the accounting ratios we ran in to significant multicollinearity issues, hence these results are not reported.

[7] Recall that all of the dynamic logit models in section three use lagged information to predict the probability of financial distress.

[8] The number of bankrupt firms differs according to the availability of data on the variables used to estimate the various models (see also the discussion in section two.).

are statistically significant predictors of the probability of bankruptcy, the ZDECOMP model only correctly predicts that around 72% of those firms that actually went bankrupt were more likely to go bankrupt than not. This is quite a deterioration in explanatory capability relative to the ZSCORE model. Things improve once we include market-driven variables in the hazard model. The model using market variables only (MV) correctly classifies around 85% of those firms that actually go bankrupt as more likely than not to go bankrupt while the model including market-driven variables and accounting ratios reflecting earnings to total assets, and book leverage (MVACC) correctly classifies around 89% of bankrupt firms as more likely than not to go bankrupt. Further, the MVACC model classifies a higher proportion of firms that go bankrupt into the first decile. Thus, while all of the models perform well in terms of correctly classifying those firms more likely to go bankrupt, the best performer by some way is the model that combines market-driven variables with accounting ratios capturing profitability and leverage.

### 5.2 Out-of-sample Performance

**INSERT TABLE 5 ABOUT HERE**

To examine the out-of-sample performance of the models in section three, we re-estimate the hazard models using data for the period 1981–1990 and then use these parameter estimates to forecast corporate financial distress over the period 1991–2006.[9] Table 5 reports the results. As in the previous subsection, deciles one through five contain firms that are predicted as more likely to go bankrupt, decile one containing those firms with the highest predicted probability of bankruptcy, while deciles six through ten contain those firms that are considered least likely to go bankrupt, decile ten containing those firms with the lowest predicted probability of bankruptcy. We then calculate the percentage of bankrupt firms that are allocated to the various deciles. Again, all of the models do reasonably well with, for example, the ZSCORE model correctly identifying 80% of those firms that go bankrupt over the period 1991–2006. Again, however, it is the model with market-driven variables and accounting ratios capturing profitability and leverage that dominates, both in terms of the percent of firms correctly classified and the percent of firms allocated to the decile containing firms with the highest probability of financial distress.

[9] We start from 1981 as our sample does not contain any bankrupt firms in 1980.

The results in this section, then, reinforce the results from section three: it is the model with market-driven variables coupled with accounting ratios capturing leverage and profitability distress, that provides the best in- and out-of-sample performance in terms of correctly classifying firms that go bankrupt.

# 6 Concluding remarks

We assess the contribution of accounting and market-based information to the prediction of the probability of financial distress for UK firms using the hazard model approach of Shumway (2001). We do this by first examining the extent to which accounting ratios alone are related to the prediction of bankruptcy. In particular, we construct the Z-score using the model in Taffler (1983) and use this to forecast the probability of bankruptcy. We also explore whether the accounting ratio components of Z-score considered individually can predict bankruptcy. We then use the market-driven variables documented in Shumway (2001) to investigate the extent to which market-related information is associated with the prediction of financial distress for UK firms. Finally, we explore the ability of accounting ratios *and* market-based variables to predict financial distress.

Our results show that in a univariate model, Z-score is a significant predictor of the probability of financial distress. This seems a reassuring result. However, when we decompose Z-score into its four accounting-related components we find that only two accounting ratios are associated with the prediction of financial distress. We also provide strong evidence that the market-based variables, i.e, market capitalization, excess returns and return volatility, significantly predict bankruptcy. Moreover, when we combine both the Z-score and the Z-score components with the market-based variables, we find that neither the Z-score nor the Z-score components contain any incremental information relevant to the prediction of financial distress above that contained in the market variables. In other words, all of the predictive content that Z-score contains in terms of the probability of financial distress is captured by excess returns, market capitalization and return volatility. This suggests that accounting information, or at least that used in calculating the Z-score, offers little extra information to that contained in market-related variables. To further explore this, we use another set of accounting ratios that Shumway (2001) and Campbell, Hilscher and Szilagyi (2008) have found significantly predict bankruptcy elsewhere. We find that in addition to market-based variables, book leverage can significantly predict financial distress. Therefore, our results suggest that an amalgam of accounting and market-based variables best captures the probability of financial distress.

In-sample and out-of sample forecasts show that the hazard model combining accounting and market-related variables and the hazard model using market variables alone have superior predictive power to the one based on the accounting-based components of Z-score. We also provide evidence that the model including both accounting and market variables outperforms the model based on market-driven variables alone.

Overall, our findings provide important insight on the prediction of corporate bankruptcy, which is of major concern to both academics and practitioners for two important reasons. First, we shed light on the role of accounting and market information on the prediction of the probability of financial distress. We find that a model combining accounting and market-based variables leads to the most powerful prediction of corporate bankruptcy. Second, we argue that Z-score needs to be treated with caution when predicting financial distress. We show that only half of the Z-score components are related to the prediction of financial distress. More importantly, Z-score and its components add no incremental information with respect to bankruptcy prediction when they are combined with market variables.

## References


Agarwal, V. and Taffler, R.J., 'Twenty-five years of the Taffler z-score model: does it really have predictive ability?', *Accounting and Business Research*, Vol. 37, 2007, pp. 285–300.

Agarwal, V. and Taffler, R.J., 'Comparing the performance of market-based and accounting-based bankruptcy prediction models', *Journal of Banking and Finance*, Vol. 32, 2008, pp. 1541–1551.

Altman, E., 'Financial ratios, discriminant analysis and the prediction of corporate bankruptcy', *Journal of Finance*, Vol. 23, 1968, pp. 589–609.

Altman, E.I., Haldeman, R.G. and Narayanan, P., 'Zeta analysis: a new model to identify bankruptcy risk of corporations', *Journal of Banking and Finance*, Vol. 1, 1977, pp. 29–54.

Beaver, W.H., 'Financial ratios as predictors of failure', *Journal of Accounting Research*, Vol. 4, 1966, pp. 71–111.

Beaver, W.H., McNichols, M.F. and Rhie, J., 'Have financial statements become more informative? Evidence from the ability of financial ratios to predict bankruptcy', *Review of Accounting Studies*, Vol. 10, 2005, pp. 93–122.

Bharath, S.T. and Shumway, T., 'Forecasting default with the Merton distance to default model', *Review of Financial Studies*, Vol. 21, 2008, pp. 1339–1369.

Byoun, S., 'How and when do firms adjust their capital structures toward targets?', *Journal of Finance*, Vol. 63, 2000, pp. 3069–3096.

Campbell, J.Y., Hilscher, T. and Szilagyi, Y., 'In search of distress risk', *Journal of Finance*, Vol. 63, 2008, pp. 2899–2939.

Dichev, I.D., 'Is the risk of bankruptcy a systematic risk?', *Journal of Finance*, Vol. 53, 1998, pp. 1131–1147.

Duffie, D., Saita, L. and Wang, K., 'Multi-Period Corporate Failure Prediction with Stochastic Covariates,' *Journal of Financial Economics*, Vol. 83, 2007, pp. 635–665.

Graham, J.R., 'How big are the tax advantages of debt?', *Journal of Finance*, Vol. 55, 2000, pp. 1901–1941.

Griffin, J.M. and Lemmon, M., 'Book-to-market equity, distress risk and stock returns', *Journal of Finance*, Vol. 57, 2002, pp. 2317–2336.

Hillegeist, S.A., Keating, E.A., Cram, D.P. and Lundstedt, K.G., 'Assessing the probability of bankruptcy', *Review of Accounting Studies*, Vol. 9, 2004, pp. 5–34.

Lau, A.H., 'A five-state financial distress prediction model', *Journal of Accounting Research*, Vol. 25, 1987, pp. 127–138.

Merton, R.C., 'On the pricing of corporate debt: the risk structure ofinterest rates', *Journal of Finance*, Vol. 29, 1974, pp. 449–470.

Ohlson, J., 'Financial ratios and the probabilistic prediction of bankruptcy', *Journal of Accounting Research*, Vol. 18, 1980, pp. 109–181.

Shumway, T., 'Forecasting bankruptcy more accurately: a simple hazard model', *Journal of Business*, Vol. 74, 2001, pp. 101–124.

Taffler, R.J., 'The assessment of company solvency and performance using a statistical model', *Accounting and Business Research*, Vol. 15, 1983, pp. 295–308.

Vassalou, M. and Xing, Y., 'Default risk in equity returns,' *Journal of Finance*, Vol. 59, 2004, pp. 831–868.

Vuong, Q.H., 'Likelihood ratio tests for model selection and non-nested hypotheses,' *Econometrica*, Vol. 57, 1989, pp. 307–333.

Zmijewski, M., 'Methodological issues related to the estimation of financial distress prediction models', *Journal of Accounting Research*, Vol. 22, 1984, pp. 59–82.

**Table 1**
**Definition of Variables**

This table defines the variables used in the study. The accounting data is from Datastream. Numbers in parentheses correspond to the Datastream code. Equity market data is taken from the London Share Price Database (LSPD).

| Variable Name | Variable definition |
|---|---|
| PROF | $\dfrac{\text{Profit before tax (384)}}{\text{Current liabilities (389)}}$ |
| Total liabilities | Total assets (392) − Equity capital & reserves (305) |
| WCAP | $\dfrac{\text{Current assets (376)}}{\text{Total liabilities}}$ |
| FRISK | $\dfrac{\text{Current liabilities (389)}}{\text{Total assets}}$ |
| Quick assets | Current assets − Total inventories (364) |
| LIQUID | $\dfrac{\text{Quick assets} - \text{current liabilities (389)}}{\left(\dfrac{\text{Sales (104)} - \text{profit before tax} - \text{depreciation (696)}}{365}\right)}$ |
| Z-score | $3.20 + 12.18 * \text{PROF} + 2.50 * \text{WCAP} - 10.68 * \text{FRISK} + 0.029 * \text{LIQUID}$ |
| EBITDA_TA | $\dfrac{\text{Earnings before interest, tax and depreciation (154+153+696)}}{\text{Total assets}}$ |
| BLEV | $\dfrac{\text{Total debt (1301)}}{\text{Total debt} + \text{Total share capital \& reserves (307)}}$ |
| REL_SIZE | $\ln\left(\dfrac{\text{Market value of equity}}{\text{Market value of FTSE all share index}}\right)$ |
| $r_{i,t}$ | stock return for firm $i$ at time $t$ (LSPD) |
| $r_{FTALL,t}$ | return on the FT All Share Index at time $t$ (LSPD) |
| EXRET | $r_{i,t-1} - r_{FTALL,t-1}$ (LSPD) |
| SIGMA | standard deviation of $\epsilon_{i,t}$ in the regression : $r_{i,t}=\alpha + \beta r_{FTALL,t} + \epsilon_{i,t}$ |

**Table 2**
**Summary Statistics**

This table presents descriptive statistics for the variables used in this study. The initial sample consists of 3,459 firms for the period 1980-2006. We identify 310 financially distressed firms and 3,149 non-financially distressed firms. The variables are winsorized at the 1% fractile in either tail of distribution, apart from relative size which is normally distributed. PROF is measured as profit before tax divided by current liabilities; WCAP is the ratio of current assets to total liabilities; FRISK is measured as current liabilities to total assets; LIQUID is defined as (quick assets minus current liabilities) divided by (sales minus profit before tax minus depreciation divided by 365); Z-score is calculated as $3.20 + 12.18\text{PROF} + 2.50\text{WCAP} - 10.68\text{FRISK} + 0.029\text{LIQUID}$. EBITDA_TA is the ratio of EBITDA to total assets. Book leverage is measured as the book value of debt divided by the book value of debt plus stockholders' equity. REL_SIZE is the natural logarithm of the firm's annual market capitalization relative to the market capitalization of the FTSE ALL SHARE index. EXRET is the firm's annual returns in excess of the return on the FTSE ALL SHARE index. SIGMA is idiosyncratic return volatility. It is estimated as the standard deviation of the residuals from a regression of each stock's monthly return on the monthly return on the FTSE ALL SHARE index.

| Variable | N | Mean | Median | Std.dev | Min | Max |
|---|---|---|---|---|---|---|
| PROF | 22,785 | -0.05 | 0.19 | 1.11 | -6.57 | 1.28 |
| WCAP | 22,785 | 1.46 | 1.05 | 1.84 | 0.09 | 13.37 |
| FRISK | 22,785 | 0.38 | 0.37 | 0.19 | 0.03 | 1.07 |
| LIQUID | 22,785 | 0.11 | -0.02 | 0.79 | -1.38 | 5.65 |
| Z-score | 22,785 | 3.01 | 3.40 | 8.63 | -18.42 | 18.42 |
| EBITDA_TA | 27,796 | 0.09 | 0.12 | 0.21 | -1.20 | 0.42 |
| Book Leverage | 27,796 | 0.26 | 0.23 | 0.24 | 0.00 | 1.44 |
| REL_SIZE | 28,503 | -2.75 | -2.92 | 2.07 | -13.22 | 4.82 |
| EXRET | 28,503 | 0.02 | 0.01 | 0.49 | -1.25 | 1.80 |
| SIGMA | 28,503 | 0.11 | 0.09 | 0.08 | 0.02 | 0.49 |

**Table 3**

**Results For Hazard Models Predicting the Probability of Financial Distress Using Different Information Sets**

This table contains results for the various hazard models we estimate. The dependent variable is an indicator variable that equals zero if the firm is not financially distressed. If the firm is financially distressed, then the dependent variable equals one only for its last firm-year observation. The independent variables are lagged to ensure that the data are observable prior to the event of financial distress. Panel A contains parameter estimates and test of their significance for the various hazard models we estimate to predict the probability of financial distress. The column headed ZSCORE contains results for a univariate hazard model that uses only Z-score to predict bankruptcy. The ZDECOMP column contains results from a hazard model where the predictor variables are the individual components of the Z-score. The MV column contains results from a hazard model that uses the market-based variables to predict financial distress. The MVZSCORE column reports results from a hazard model that combines the Z-score with the market-based predictor variables. The MVZDECOMP column reports the results derived from a hazard model that combines the individual components of the Z-score with the market-based predictors. while the MVACC column contains results from a hazard model using market-based predictors and accounting ratios measuring earnings to total assets, and book leverage. Finally, the MVACCZ column shows the results from a hazard model that combines the Z-score with both the earnings and book leverage accounting variables, and the market-based variables. The row labeled Wald Statistic contains the Wald test testing the hypothesis that the coefficients are jointly zero. It is distributed as $\chi^2(k)$, where $k$ is the number of parameters (excluding the constant.) We scale the Wald-Chi Square statistic by the average number of observations per firm. Figures in parentheses are scaled Wald statistics testing the hypothesis that the individual coefficient is zero. These have a $\chi^2(1)$ distribution. Panel B contains the results from Vuong tests for model comparison. Under the null hypothesis that there is no difference between the two models, the log of the ratio of the likelihood for model $i$ to that for model $j$ should be zero. If the difference is significantly positive, $i$ is preferred to $j$ and vice versa. $***$, $**$ and $*$ denote significance at the 1, 5 and 10 percent levels respectively.

Panel A: Bankruptcy Prediction Models For UK Firms

| | ZSCORE | ZDECOMP | MV | MVZSCORE | MVZDECOMP | MVACC | MVACCZSCORE |
|---|---|---|---|---|---|---|---|
| Constant | −4.7329*** | −5.6872*** | −6.3067*** | −6.2375*** | −6.8736*** | −6.4965*** | −6.5547*** |
| | (479.80) | (110.31) | (179.23) | (119.46) | (101.83) | (149.25) | (113.67) |
| PROF | | −0.3355*** | | | −0.3332 | | |
| | | (7.21) | | | (0.04) | | |
| WCAP | | −0.0206 | | | −0.0004 | | |
| | | (0.02) | | | (0.00) | | |
| FRISK | | 2.1191** | | | 1.1247 | | |
| | | (5.52) | | | (1.36) | | |
| LIQUID | | 0.0003 | | | 0.0003 | | |
| | | (0.00) | | | (0.13) | | |
| Z-score | −0.0795*** | | | −0.0305 | | | −0.0305 |
| | (15.72) | | | (1.61) | | | (0.87) |
| EBITDA_TA | | | | | | −0.1210 | −0.4021 |
| | | | | | | (0.03) | (0.22) |
| BLEV | | | | | | 1.1238** | 0.7553 |
| | | | | | | (5.00) | (1.66) |
| REL_SIZE | | | −0.2144* | −0.2065 | −0.2189 | −0.2035* | −0.2203* |
| | | | (3.50) | (2.57) | (2.82) | (3.03) | (2.89) |
| EXRET | | | −0.9856*** | −0.7802** | −0.8668** | −0.8569** | −0.7952** |
| | | | (9.32) | (4.66) | (5.71) | (6.43) | (4.69) |
| SIGMA | | | 5.9369*** | 5.1522** | 5.7623** | 5.1454** | 5.2495** |
| | | | (9.97) | (5.47) | (6.94) | (6.18) | (5.41) |
| Wald statistic | 127.62*** | 107.91*** | 291.43*** | 262.69*** | 261.82*** | 352.25*** | 284.48*** |
| Number of observations | 22,785 | 22,785 | 28,503 | 21,982 | 21,982 | 27,796 | 21,964 |

Panel B: Vuong Tests

| Model $i$ versus Model $j$ | $z$ statistic |
|---|---|
| MV versus ZSCORE | 5.25*** |
| MV versus ZDECOMP | 6.86*** |
| MVACC versus ZSCORE | 6.80*** |
| MVACC versus MV | 3.70*** |
| MVACC versus MVACCZSCORE | 3.92*** |

**Table 4**
**In-sample Forecast Accuracy Examining The Percentage of Firms Predicted To Go Bankrupt That Actually Went Bankrupt**

This table examines the in-sample forecast accuracy of four of the hazard models we estimate. Firms are sorted in to deciles based on their estimated probability of financial distress. Decile 1 contains those firms with the highest probability while decile 10 contains those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt the models place in to each decile. The column headed ZSCORE contains results for a univariate hazard model that uses only Z-score to predict bankruptcy. The ZDECOMP column contains results from a hazard model where the predictor variables are the individual components of the Z-score. The MV column contains results from a hazard model that uses the market-based variables to predict financial distress while the MVACC column contains results from a hazard model using market-based predictors and accounting ratios measuring earnings to total assets, and book leverage.

| Decile | ZSCORE | ZDECOMP | MV | MVACC |
|---|---|---|---|---|
| 1 | 30.50 | 30.54 | 36.43 | 44.49 |
| 2 | 18.22 | 14.77 | 18.96 | 13.69 |
| 3 | 18.22 | 11.03 | 11.90 | 11.90 |
| 4 | 7.59 | 8.87 | 11.15 | 14.44 |
| 5 | 5.91 | 7.39 | 7.06 | 4.94 |
| 6-10 | 19.70 | 27.59 | 14.50 | 11.41 |
| No. of Bankrupt Firms | 203 | 203 | 269 | 263 |

**Table 5**
**Out-of-sample Forecast Accuracy Examining The Percentage of Firms Predicted To Go Bankrupt That Actually Went Bankrupt**

This table examines the out-of-sample forecast accuracy of four of the hazard models we estimate. The models are estimated using data over the period 1981–1990. These parameter estimates are then used to calculate the probability of financial distress over the period 1991–2006. Firms are sorted in to deciles based on their estimated probability of financial distress. Decile 1 contains those firms with the highest probability while decile 10 contains those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt the models place in to each decile. The column headed ZSCORE contains results for a univariate hazard model that uses only Z-score to predict bankruptcy. The ZDECOMP column contains results from a hazard model where the predictor variables are the individual components of the Z-score. The MV column contains results from a hazard model that uses the market-based variables to predict financial distress while the MVACC column contains results from a hazard model using market-based predictors and accounting ratios measuring earnings to total assets, and book leverage.

| Decile | ZSCORE | ZDECOMP | MV | MVACC |
|---|---|---|---|---|
| 1 | 27.86 | 28.85 | 35.58 | 36.72 |
| 2 | 20.90 | 17.91 | 17.79 | 20.29 |
| 3 | 17.91 | 10.94 | 15.38 | 15.46 |
| 4 | 7.46 | 8.46 | 9.13 | 9.66 |
| 5 | 5.97 | 6.97 | 6.25 | 6.75 |
| 6-10 | 19.90 | 26.87 | 15.87 | 11.11 |
| No. of Bankrupt Firms | 201 | 201 | 208 | 207 |

# Appendix

This appendix contains results from estimating the various hazard models with a common dataset.

**Table A1**
**Results For Hazard Models Predicting the Probability of Financial Distress Using a Balanced Panel**

This table contains results for the various hazard models we estimate, using a balanced panel. The dependent variable is an indicator variable that equals zero if the firm is not financially distressed. If the firm is financially distressed, then the dependent variable equals one only for its last firm-year observation. The independent variables are lagged to ensure that the data are observable prior to the event of financial distress. Panel A contains parameter estimates and test of their significance for the various hazard models we estimate to predict the probability of financial distress. The column headed ZSCORE contains results for a univariate hazard model that uses only Z-score to predict bankruptcy. The ZDECOMP column contains results from a hazard model where the predictor variables are the individual components of the Z-score. The MV column contains results from a hazard model that uses the market-based variables to predict financial distress. The MVZSCORE column reports results from a hazard model that combines the Z-score with the market-based predictor variables. The MVZDECOMP column reports the results derived from a hazard model that combines the individual components of the Z-score with the market-based predictors. while the MVACC column contains results from a hazard model using market-based predictors and accounting ratios measuring earnings to total assets, and book leverage. Finally, the MVACCZ column shows the results from a hazard model that combines the Z-score with both the earnings and book leverage accounting variables, and the market-based variables.The row labeled Wald Statistic contains the Wald test testing the hypothesis that the coefficients are jointly zero. It is distributed as $\chi^2(k)$, where $k$ is the number of parameters (excluding the constant.) We scale the Wald-Chi Square statistic by the average number of observations per firm. Figures in parentheses are scaled Wald statistics testing the hypothesis that the individual coefficient is zero. These have a $\chi^2(1)$ distribution. Panel B contains the results from Vuong tests for model comparison. Under the null hypothesis that there is no difference between the two models, the log of the ratio of the likelihood for model $i$ to that for model $j$ should be zero. If the difference is significantly positive, $i$ is preferred to $j$ and vice versa. $***$, $**$ and $*$ denote significance at the 1, 5 and 10 percent levels respectively.

Panel A: Bankruptcy Prediction Models For UK Firms

| | ZSCORE | ZDECOMP | MV | MVZSCORE | MVZDECOMP | MVACC | MVACCZ |
|---|---|---|---|---|---|---|---|
| Constant | −4.7242*** | −5.7890*** | −6.5379*** | −6.2421*** | −6.8779*** | −6.7176*** | −6.5547*** |
| | (440.62) | (102.21) | (147.77) | (119.21) | (101.76) | (126.38) | (113.67) |
| PROF | | −0.3528*** | | | −0.0345 | | |
| | | (7.40) | | | (0.05) | | |
| WCAP | | −0.0248 | | | −0.0010 | | |
| | | (0.03) | | | (0.00) | | |
| FRISK | | 2.3659** | | | 1.1272 | | |
| | | (6.13) | | | (1.37) | | |
| LIQUID | | 0.0005 | | | 0.0003 | | |
| | | (0.20) | | | (0.13) | | |
| Z-score | −0.0834*** | | | −0.0306 | | | −0.0305 |
| | (15.59) | | | (1.62) | | | (0.87) |
| EBITDA_TA | | | | | | −0.0730 | −0.4021 |
| | | | | | | (0.01) | (0.22) |
| BLEV | | | | | | 0.9770* | 0.7553 |
| | | | | | | (3.23) | (1.66) |
| REL_SIZE | | | −0.2436* | −0.2075 | −0.2200 * | −0.2333* | −0.2203* |
| | | | (3.69) | (2.59) | (2.84) | (3.28) | (2.89) |
| EXRET | | | −0.9143*** | −0.7848** | −0.8714** | −0.8328 ** | −0.7952** |
| | | | (6.79) | (4.70) | (5.76) | (5.14) | (4.69) |
| SIGMA | | | 6.1998*** | 5.1647** | 5.7665*** | 5.5908 ** | 5.2495 ** |
| | | | (9.39) | (5.50) | (6.95) | (6.35) | (5.41) |
| Wald statistic | 132.54*** | 120.52*** | 242.60*** | 263.22*** | 262.35*** | 282.25*** | 284.48 *** |
| Number of observations | 21,964 | 21,964 | 21,964 | 21,964 | 21,964 | 21,964 | 21,964 |

Panel B: Vuong Tests

| Model $i$ versus Model $j$ | $z$ statistic |
|---|---|
| MV versus ZSCORE | 3.84*** |
| MV versus ZDECOMP | 5.22*** |
| MVACC versus ZSCORE | 5.06*** |
| MVACC versus MV | 2.69** |
| MVACC versus MVACCZSCORE | −1.23 |