


On corporate financial distress prediction: What can we learn from private firms in a developing economy? Evidence from Greece

Evangelos C. Charalambakis¹ · Ian Garrett² 

Published online: 19 March 2018
© The Author(s) 2018

Abstract Using a large dataset that includes nearly 31,000 Greek private firms we examine the determinants of the probability of corporate financial distress. Using a multi-period logit model, we find that profitability, leverage, the ratio of retained earnings-to-total assets, size, the liquidity ratio, an export dummy variable, the tendency to pay out dividends and the growth rate in real GDP are strong predictors of the probability of financial distress for Greek private firms. A model including these variables exhibits the highest in-sample and out-of-sample performance in terms of correctly classifying firms that went bankrupt as more likely to go bankrupt. The predictive ability of the model remains when we increase the forecast horizon, suggesting that the model works well over short and longer time horizons.

Keywords Corporate financial distress · Bankruptcy prediction · Private firms · Discrete hazard (multi-period logit) model · Financial statements · Greek debt crisis

JEL Classification G13 · G17 · G33 · C41

1 Introduction

The failure of a firm is an event of major concern in economic life. Consequently, the prediction of corporate financial distress has received considerable attention in the field of corporate finance. Academic researchers and practitioners have developed various models to forecast financial distress using either accounting or stock market information. Following the seminal work of Beaver (1966) and Altman (1968), there have been a

✉ Ian Garrett
ian.garrett@manchester.ac.uk

¹ Economic Research Department, Bank of Greece, 21 E. Venizelos Avenue, 102 50 Athens, Greece

² Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK

substantial number of studies that make use of accounting ratios to predict corporate bankruptcy.¹ More recent studies have focused on Merton's (1974) structural market-based model for pricing corporate debt to derive the probability of corporate default.² Shumway (2001) and Campbell et al. (2008) use a discrete hazard model that incorporates both accounting and stock-market-based variables to forecast bankruptcy.

Despite the fact that predicting the probability of corporate financial distress has received a significant amount of attention, our understanding of the prediction of corporate financial distress is far from complete. There are two gaps in the literature that this paper goes some way towards filling. Much of the literature to-date has focused on publicly-traded firms in developed economies, for which data is more widely available. Much less attention has been devoted to the determinants of the probability of corporate financial distress for private firms and less still to the determinants of the probability of financial distress of private firms in developing economies. Indeed, to our knowledge there is no empirical research that addresses the question of what drives the likelihood of private firms entering financial distress in a developing economy.³ We attribute this to the difficulty of obtaining private firms' financial data and information on when a private firm goes bankrupt.

This paper contributes to the literature by examining the determinants of the probability of financial distress for private firms in a developing economy, that of Greece. There are several reasons why examining the determinants of the probability of financial distress for private firms is important. First, private firms, especially small- and medium-sized enterprises (SMEs) are of key importance for economic activity, employment and innovation in many countries. With respect to Greece, for example, 99.9% of Greek firms are defined as SMEs, they account for 84.9% of the total workforce and 69% of the value added in the economy (OECD 2015). Similarly, SMEs are the backbone of the economy in the Eurozone. According to Wymenga et al. (2012), across the Eurozone SMEs account for around 98% of all Euro area firms, around 75% of total employment and generate around 60% of gross value added while over 99% of US and UK firms are privately owned and are responsible for more than half of the GDP of the US and the UK, respectively. Second, the evidence shows that private firms are different from public ones. Private firms are smaller, more leveraged, more dependent on trade credit and bank loans, invest more, and face higher borrowing costs.⁴ These differences between private and public firms raise the issue of whether the (non-stock-market) factors that predict the probability of bankruptcy for public firms are (a) the same for private firms and (b) have the same sign as those for public firms and perhaps similar orders of magnitude as well. Third, providing information on the probability of bankruptcy for private firms is quite challenging as these firms do not have their shares traded on a stock exchange and hence there is a lack of market data. As a result, the information we can use to predict bankruptcy is mainly derived from financial statements. Therefore, examining whether it is possible to successfully predict the probability of financial distress for private firms is essential, not least because it will help in

¹ See, among many others, Ohlson (1980), and Zmijewski (1984).

² See Vassalou and Xing (2004), Duffie et al. (2007) and Bharath and Shumway (2008), among others.

³ There have been some papers in the more recent literature that focus on the probability of financial distress for publicly traded firms in developing economies. Kwak et al. (2012) examine the probability of financial distress for Korean firms while Charalambakis and Garrett (2016) use discrete hazard analysis to investigate the determinants of financial distress for Indian firms.

⁴ See, for example, Petersen and Rajan (1994), Brav (2009), Asker et al. (2012), Gao et al. (2012), Michaely and Roberts (2012), and Faccio et al. (2012).

formulating policy relating to the supply of credit by banks and the cost of credit for this type of firm. Evidence on the performance of bankruptcy forecast models for private firms will also provide insights into the usefulness of information contained in financial statements for private firms.

Using a large dataset of Greek private companies, we estimate the probability of corporate financial distress for Greek private firms over the period 2003–2011. In addition to being a developing economy where private firms predominate, Greece provides a challenging test for models predicting the probability of bankruptcy given the debt crisis that enveloped the country from 2009 to 2011 following the global financial crisis. Further, examining bankruptcy around the Greek debt crisis allows us to see the degree to which business failures in the crisis period could be predicted by pre-crisis data.

We develop multi-period logit (discrete hazard) models to evaluate the variables that are significantly associated with the prediction of financial distress for Greek private firms.⁵ In the literature to date, discrete hazard models have tended to be used to predict the probability of financial distress for publicly-traded firms in developed economies and tend to use both accounting information, usually in the form of accounting ratios, and stock market information; see, for example, Shumway (2001), Campbell et al. (2008) and Charalambakis and Garrett (2016). Studies making use of accounting information only have tended to use discriminant analysis or, more predominantly, traditional logit or probit models; see, for example, Altman (1968), Ohlson (1980), Taffler (1983) and Zmijewski (1984). These latter models are single-period (static) classification models in which the probability of bankruptcy does not change over time as the models ignore the fact that firms change through time. Shumway (2001) shows that estimating the probability of financial distress using static models leads to biased and inconsistent estimates of the probability of bankruptcy and incorrect inferences concerning the statistical significance of the predictor variables. Shumway (2001) proposes estimating the probability of bankruptcy using a discrete hazard model which is a dynamic model that captures the fact that firms change over time, enabling the probability of bankruptcy to change over time as a function of a vector of explanatory variables that also vary with time. Shumway (2001) finds that using a discrete hazard model to estimate the probability of bankruptcy, up to half of the accounting ratios that had previously been used to forecast bankruptcy for US firms are not statistically significant, that is, they are not statistically related to failure. Given this and given that there seems no reason why the probability of financial distress does not change over time for private firms, the discrete hazard model or, equivalently, the multi-period logit model, seems an appropriate choice of model to estimate the probability of financial distress, notwithstanding the presence of a financial crisis in our case, which one might expect to cause a firm's probability of financial distress to change.

We find strong evidence that a model that contains seven firm-specific variables, along with the growth rate of real GDP and a dummy variable that captures the Greek debt crisis best explains the likelihood that a Greek private firm will experience financial distress. In particular, we find that profitability, retained earnings-to-total assets, size and liquidity are negatively associated with the probability of financial distress for Greek private firms while leverage is positively related to the probability of financial distress. We also find strong evidence that the probability of a private firm going bankrupt decreases when it exports and

⁵ Shumway (2001) shows that a discrete hazard model is equivalent to a multi-period logit model where the multi-period logit model is defined as, "...a logit model that is estimated with data on each firm in each year of its existence as if each firm year were an independent observation." (Shumway 2001, p. 105). We use the terms discrete hazard and multi-period logit interchangeably.

pays out dividends. Interestingly, we find that the dummy variable for the Greek debt crisis is negatively associated with the probability of financial distress, meaning that bankruptcies were less common during the crisis than might be expected. This finding is borne out by the average number of bankrupt firms within the crisis period decreasing to more than half those in the pre-crisis period. In particular, the average number of bankrupt firms was 271 for the period 2006–2008 while the average number of bankrupt firms was 134 for the crisis period of 2009–2011. In terms of macroeconomic influences on the probability of financial distress we find that the growth rate in real GDP is negatively related to the probability of financial distress. Forecast accuracy tests show that in-sample, this model correctly classifies 87% of those firms that subsequently went bankrupt as likely to go bankrupt. In addition, the results do not change when we extend the forecast horizon from one year to two and three years, suggesting that the model is quite robust with respect to the horizon over which the probability of financial distress is predicted, at least in the short- to medium-term.

To examine the out-of-sample forecasting performance of our preferred model, and to examine the extent to which our findings may be driven by the debt crisis that enveloped Greece following the global financial crisis, we re-estimate our preferred model excluding the Greek debt crisis to investigate whether the nature of the relationship between the predictor variables and the probability of bankruptcy changes and to examine whether, out-of-sample, our preferred model is still able to correctly classify those firms that subsequently went bankrupt as likely to go bankrupt. With the exception of size, which becomes insignificant, the estimated parameters have the same sign and are of similar magnitude to the full-sample estimates. More importantly, the model correctly classifies 88% of firms that went bankrupt during the Greek crisis as likely to go bankrupt.

We also investigate whether the results differ when we split the sample into small and medium-sized private firms based on the number of employees.⁶ We find that the signs and the magnitude of the estimated coefficients of the seven firm-specific variables vary across small and medium firms.

The rest of the paper is organized as follows. Section 2 provides a review of the literature and some methodological background to modelling the probability of financial distress using the discrete hazard/multi-period logit approach. Section 3 describes the dataset. Section 4 presents the results for the various models we estimate, the results from forecast accuracy tests and robustness checks. Section 5 concludes.

2 Literature review and empirical design

Several econometric techniques have been used to predict financial distress for publicly traded firms. Beaver (1966) uses univariate analysis to investigate the ability of accounting data to predict bankruptcy. Based on this method a financial ratio for the firm of interest is compared to a benchmark ratio to discriminate a failed firm from a non-failed firm. Altman (1968) employs multiple discriminant analysis to determine the Z-score, a widely used measure for predicting bankruptcy. The objective of discriminant analysis is to generate a linear combination of variables that best separate the bankrupt firms from the non-bankrupt ones. Although popular, this method is subject to some limitations. First, it fails to provide a test of the significance of the individual variables. Second, it assumes that the predictors

⁶ One of the ways The European Union defines firm size is based on the number of employees (Commission 2003).

are distributed as multivariate normal. Third, it prevents the use of dummy variables that can enhance the predictive ability of the bankruptcy forecast models.

Altman et al. (1977) use quadratic discriminant analysis to forecast bankruptcy. 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. Another strand of the literature draws on the Merton (1974) option pricing framework to derive the probability of default. In these models, equity is viewed as a call option on the firm’s assets and the probability of going bankrupt is simply the probability that the call option is worthless at maturity, i.e., the market value of total assets is less than the face value of total liabilities. Vassalou and Xing (2004) employ the Merton model to investigate whether default risk is priced in equity returns. Duffie et al. (2007) show that the default probabilities derived from Merton’s model can have significant predictive power and consequently can generate a term structure of default probabilities.

Etheridge and Sriram (1997) examine the performance of neural networks with respect to corporate financial distress prediction. They find that the neural network outperforms multivariate discriminant analysis and logistic models. Nittayagasetwat (1996) applies a neural network to forecast bankruptcy for US firms. He finds that his neural network model exhibits higher predictive ability than a logit model. While neural network techniques can provide higher classification rates, they cannot provide any information on the significance of the predictor variables. It is therefore more difficult to assess the contribution of the predictor variables to the prediction of financial distress.

Shumway (2001) argues and shows that the models used by Altman (1968), Ohlson (1980) and Zmijewski (1984) are misspecified as they do not properly address the length of time that a healthy firm has survived, thereby inducing bias. He overcomes this problem by employing a discrete hazard model which is econometrically equivalent to a multi-period logit model. This model has two main advantages. First, it allows researchers to take advantage of all the available firm-year observations. Second, it is a dynamic model in the sense that it enables the probability of bankruptcy to change over time as a function of a vector of explanatory variables that also vary with time. In his empirical work, Shumway (2001) finds that a discrete hazard model delivers the best performance in terms of out-of-sample predictive ability. Since the seminal work of Shumway (2001), Hillegeist et al. (2004) and Agarwal and Taffler (2008) have used discrete hazard models to compare the performance of accounting-information-based and stock-market-information-based bankruptcy prediction models for US and UK firms, respectively. Chava and Jarrow (2004) extend the model of Shumway (2001), highlighting the importance of including industry effects in the discrete hazard model. They also provide evidence that the predictive power of accounting variables weakens when stock-market-based variables are included in the model. Bharath and Shumway (2008) use a discrete hazard model to examine the contribution of the probability of default derived from the Merton (1974) model to predicting financial distress. They find that the probability of default based on the Merton model is not a sufficient statistic for default. Using US data, Campbell et al. (2008), building on the work of Shumway (2001), find that the default probability based on the Merton model has relatively little forecasting power in a discrete hazard model when conditioning on accounting and stock-market-based variables. Tinoco and Wilson (2013) compare hazard models with neural networks and Z-score models using UK data. They find that a panel logistic model with time-varying covariates, which is equivalent to a hazard model, that combines accounting, stock market and macroeconomic variables outperforms the neural network approach and Z-score approach. Similarly, Bauer and Agarwal (2014) show that

for the UK, the hazard model developed by Shumway (2001) outperforms the Z-score model developed by Taffler (1983) and the contingent claims model of Bharath and Shumway (2008).

While there is a wealth of evidence on the prediction of financial distress for publicly-listed firms in developed markets, evidence is a little harder to come by for developing markets and for private firms. Kwak et al. (2012) predict bankruptcy for Korean firms after the 1997 financial crisis. Using the accounting variables that Altman (1968) and Ohlson (1980) use, Kwak et al. (2012) estimate their bankruptcy prediction model using multiple criteria linear programming (MLCP). They find that their model works as well as traditional multiple discriminant analysis and traditional logit analysis. However, other than noting where there is a significant difference in the means of the variables for bankrupt and non-bankrupt firms, Kwak et al. (2012) do not focus on the sign and significance of the variables in the prediction of bankruptcy as they are concerned with a comparison of MLCP with traditional multiple discriminant analysis and traditional logit analysis. In examining the predictors of financial distress in the trading and services sectors in Malaysia, Alifiah (2014) finds that the likelihood of distress is negatively related to the debt ratio, the working capital ratio and net income to total assets while it is positively related to the total asset turnover ratio and the base lending rate. Charalambakis and Garrett (2016) evaluate the contribution of accounting and stock-market-based information to the prediction of bankruptcy across developed and emerging markets using discrete hazard models. While they find that book leverage combined with three stock-market-based variables best predict the probability of financial distress for UK firms, they find that stock market information has no significant role to play in predicting the probability of distress for Indian firms. Rather, they find the probability of financial distress in India is negatively related to an accounting-based measure of profitability and positively related to the ratio of current liabilities to total assets.

In terms of evidence regarding the determinants of the probability of financial distress for private firms, again it is quite sparse and almost exclusively focused on developed economies. Of those studies that investigate the probability of financial distress for private firms, Falkenstein et al. (2000) evaluate the credit risk of US private firms using Moody's RiskCalc^{TM7} and find that the relationship between financial variables and credit risk differs quite substantially across public and private firms. Cangemi et al. (2003) use Standard and Poor's Credit Risk Tracker to examine the default risk of French private firms using the maximum expected utility (MEU) approach. Altman and Sabato (2007) find that the probability of *not* defaulting for 2000 SMEs in the US is positively related to EBITDA-to-total assets, retained earnings-to-total assets, cash-to-total assets and EBITDA-to-interest expense while it is negatively related to short-term debt-book equity. Altman et al. (2010) combine both qualitative and financial information in a default prediction model for UK SMEs. In addition to finding that the accounting ratios used by Altman and Sabato (2007) significantly predict the probability of default for UK SMEs, they find that data relating to legal action by creditors to recover unpaid debts, company filings and audit reports/opinions significantly increase the performance of their model. However, such information is not always available in advance or may not be updated frequently enough to facilitate accurate predictions. Diekes et al. (2013) use a probit model and find that business credit information improves the accuracy of default predictions for private firms in Germany.

⁷ The ten ratios that are used as inputs in Moody's RiskCalcTM are: assets-to-consumer price index; inventories-to-cost of goods sold; liabilities-to-assets; net income growth; net income-to-assets; the quick ratio; retained earnings-to-assets; sales growth; cash-to-assets and the debt service coverage ratio.

Given the comparative advantages of the discrete hazard model outlined earlier compared to the models of Altman (1968), Ohlson (1980) and Zmijewski (1984), particularly that the hazard model allows the researcher to take advantage of *all* firm year observations rather than each firm having one observation (one or zero depending on whether or not it went bankrupt) we use the discrete hazard model as the basis of our empirical analysis. The model is of the following form:

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

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

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

where Y_{it} is a variable that equals one if firm i enters financial distress 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 prior to the occurrence of bankruptcy.

With regard to inference in relation to (2), Shumway (2001) shows that it is necessary to adjust the Wald statistics that test the significance of the coefficients. The reason for this stems from the fact that the discrete hazard model can be thought of as a multi-period logit model. The multi-period logit model in turn is a logit model in which each firm-year observation is treated as if it were an independent, separate firm. However, in the discrete hazard/multi-period logit model firm-year observations are not independent of each other: a firm that went bankrupt last year, for example, cannot go bankrupt this year and therefore the test statistics need to be adjusted to reflect this. Shumway (2001) shows that this is easily done through the Wald statistic: for the multi-period logit model, the test statistics need to be scaled by the average number of firm years per firm.

3 Sample and data

The sample consists of Greek private firms that operated in Greece over the period 2003–2011. The source of our data is the ICAP database, which contains annual publicly available data mainly derived from financial statements for nearly 60,000 Greek private firms.⁸ We exclude financial firms and state-owned enterprises from the

⁸ The ICAP database is the largest business database in Southeastern Europe. It is provided by the ICAP group (<http://www.icap.gr>), who are an Information Partner with Bureau van Dijk, the providers of the Amadeus database, and who have a partnership with Dun and Bradstreet. The ICAP Group is the only company in Greece recognized by the Bank of Greece as an External Credit Assessment Institution and in July 2011, ICAP was accepted by the European Securities and Markets Authority as a Credit Rating Agency.

sample.⁹ We also exclude firm-year observations for which we do not have available data. Our final sample includes 30,886 firms and 188,065 yearly observations. To forecast corporate financial distress we need to define which firms enter financial distress and when this occurs. The ICAP database identifies the year of the private firm's death and the type of death (bankruptcy, liquidation, inactivity due to insolvency state, mergers and acquisitions and resolution). We consider a private firm to be financially distressed if it has either (a) been declared bankrupt, liquidated or dissolved or (b) been inactive due to being in a state of insolvency. Firms that have managed to resolve the insolvency state or have been the subject of a merger or acquisition are regarded as non-bankrupt firms. Using these criteria for classifying bankruptcy there are 1770 bankrupt firms in our sample with 5957 firm-year observations and 29,116 non-bankrupt firms with 182,108 firm-year observations.

To explore which factors are related to the probability of financial distress for private firms we focus on ratios derived from financial statements as there is no market data available for these firms. To this end, we examine whether the accounting ratio components of the Z-score developed by Altman (1968) for the US and Taffler (1983) for the UK that are widely used in the literature can help in predicting the probability of financial distress for private firms in a developing economy. In particular, we examine the impact of profitability, leverage, retained earnings-to-total assets, size and the liquidity ratio on the probability of financial distress. We also include a dummy variable for export activity and a dividend payout dummy variable as additional firm-specific predictors of the probability of financial distress. The motivation for including a dummy variable to denote export activity comes from evidence that across a range of countries and industries, firms that export have been shown to be larger, more skill- and capital-intensive and to pay higher wages than firms that do not export. Additionally, there is evidence that good firms become exporters and that in low-income countries exporting raises productivity (see, among others, Bernard and Jensen 1995, 1999; Bernard et al. 2007; Van Biesebroeck 2005). Taken together, this evidence is suggestive of exporting firms being less likely to enter financial distress. The motivation for including a dummy variable to denote dividend payout stems from the dividend smoothing and dividend signalling literatures. In his seminal paper on the dividend setting behavior of managers, Lintner (1956) argues and finds empirical evidence to support the notion that managers are reluctant to make dividend changes that will have to be reversed at a later date, that is, managers smooth dividends in order to avoid having to reverse dividend changes because that will be costly (see also Marsh and Merton 1987; Garrett and Priestley 2000). Garrett and Priestley (2000) find evidence that in a model based on Lintner (1956) dividends will increase in response to increases in earnings, suggesting that firms will increase dividends to signal that they believe the increase is sustainable. The dividend signaling models of Bhattacharya (1979) and Miller and Rock (1985) suggest that managers use dividends as a signal to indicate the quality of their firm's future earnings. High quality firms are considered to be more capable of bearing the cost of the dividend than low-quality firms. Taken together, this suggests that a firm that pays out dividends will have a lower probability of financial distress. Table 1 provides more detail on the definition of the firm-specific variables.

⁹ The financial firms we exclude from the sample are those that have NACE codes 64-66; the NACE code industry classification is the European standard industry classification system. The NACE codes follow the NACE Revision 2 classification. The list of Greek state owned enterprises includes public utilities (water, electricity and gas) and public enterprises from other sectors such as transportation, telecommunications and the defense industry. The list of Greek state owned companies is available from the ICAP database.

Table 1 Definition of firm-specific variables

Variable	Definition
EBITDA_TA	$\frac{\text{Earnings before interest, tax and depreciation}}{\text{Total assets}}$
BLEV	$\frac{\text{Book value of total debt}}{\text{Total assets}}$
RET_TA	$\frac{\text{Retained earnings}}{\text{Total assets}}$
SIZE	Natural logarithm of total assets
LIQUID	$\frac{(\text{Current assets} - \text{current liabilities})}{\text{Total assets}}$
EXPORT	Dummy variable equal to one if a firm exports, zero otherwise
DIVPAY	Dummy variable equal to one if a firm pays out dividends, zero otherwise

To account for any industry effects that may have an impact on the prediction of financial distress we include industry dummies in all of our models. To define the industry dummies, we classify firms into industries according to the NACE industry classification code (following NACE Revision 2). The NACE industry classification code is the European standard industry classification system. Based on the two-digit NACE code classification, we place firms into 16 industry groups, constructing 16 corresponding industry dummies. The industries are: Agriculture, forestry and fishing; Mining and quarrying; Manufacturing; Construction; Electricity, gas, steam and air conditioning supply; Water supply, sewerage, waste management and remediation activities; Wholesale and retail trade; Transportation and storage; Accommodation and food service activities; Information and communication; Real estate activities; Professional, scientific and technical activities; Administrative and support service activities; Education; Human health and social work activities; and Arts, entertainment and recreation.

Finally, we include the growth rate of real GDP as a predictor variable and we include a dummy variable to capture and control for any effect the Greek debt crisis that began in 2009, following on from the global financial crisis, might have. The Greek crisis dummy variable takes the value 1 for the period 2009–2011 and zero for the pre-crisis period of 2003–2008.

To deal with extreme observations, we winsorize the independent variables at the 1st and 99th percentiles of the distribution. Descriptive statistics for the firm-specific variables for the full sample are reported in Table 2. Looking at Table 2, we observe that the profitability and debt ratios are positively skewed since the respective means are greater than the respective medians, whereas the ratio of retained earnings-to-total assets and the liquidity ratio are negatively skewed. Further, on average, private firms have negative retained earnings. Focusing on the export and dividend dummy variables, less than 25% of the firms in the sample are involved in export activity while 27% of them pay out dividends. Table 3 reports the descriptive statistics for the sample of bankrupt and non-bankrupt firms. As we would expect, bankrupt firms are on average less profitable, more levered, smaller and accumulate more losses relative to non-bankrupt firms. Bankrupt firms are also less liquid, less likely to export and less likely to pay out dividends relative to non-bankrupt firms. Panel C of Table 3 documents the difference in the means of the variables between non-bankrupt and bankrupt firms. In all cases, the null hypothesis that there is no difference between the means for bankrupt and non-bankrupt firms is rejected. The

Table 2 Descriptive statistics, full sample

Variable	Mean	Median	Standard Deviation	Min.	Max.
EBITDA_TA	0.06	0.05	0.14	- 0.52	0.56
BLEV	0.30	0.18	0.33	0.00	1.22
RET_TA	- 0.14	0.00	0.47	- 3.38	0.38
SIZE	14.12	14.09	1.49	5.03	18.41
LIQUID	- 0.00	0.00	0.36	- 1.19	0.97
EXPORT	0.23	0.00	0.42	0.00	1.00
DIVPAY	0.27	0.00	0.45	0.00	1.00

This table presents the mean, median, standard deviation, minimum and maximum values for the variables based on the entire sample of private firms. EBITDA_TA is measured as earnings before interest, tax and depreciation divided by total assets. BLEV is measured as the book value of total debt to total assets. RET_TA is the ratio of retained earnings-to-total assets. SIZE is measured as the natural logarithm of total assets. LIQUID is defined as current assets minus current liabilities, all divided by total assets. EXPORT is a dummy variable that takes the value one if the firm exports, zero otherwise. DIVPAY is a dummy variable which equals to one if a firm pays out dividends, zero otherwise

significant differences in the means of the variables suggest that bankrupt and non-bankrupt firms differ markedly and that these differences may have some ability in predicting the probability of financial distress. We turn our attention to this in the next section.

4 Results

4.1 The predictive ability of multi-period logit models

We estimate the probability of financial distress for Greek private firms using a series of multi-period logit models. The results are presented in Panel A of Table 4. We begin with a bankruptcy prediction model that incorporates three accounting ratios: profitability, leverage, and retained earnings scaled by total assets. We supplement these ratios with a measure of firm size to control for the fact that SMEs are smaller. The results for this model can be found in the column labelled BPM1. The results show that all four variables are strongly associated with the probability of financial distress for Greek private firms, with profitability, retained earnings-to-total assets and size having a negative impact on the probability of financial distress while leverage has a positive effect on the probability of financial distress. The results in column BPM1 have the expected signs and are consistent with those documented in the literature for listed firms in both developed and emerging markets and for private firms in developed markets. The negative relationship between profitability and the probability of financial distress and the positive relationship between leverage and the probability of financial distress has been well-documented for the US, for example (see, among others, Ohlson 1980; Shumway 2001; Campbell et al. 2008) while Charalambakis and Garrett (2016) find similar evidence for listed firms in India. Altman and Sabato (2007) and Altman et al. (2010) find a negative relationship between profitability and the probability of financial distress and a positive one between leverage and financial distress for US and UK SMEs respectively. Altman and Sabato (2007) and Altman et al. (2010) also document a significant negative relationship between retained

Table 3 Descriptive statistics for non-bankrupt and bankrupt firms

Variable	Mean	Median	Standard Deviation				
Panel A: Non-bankrupt firms							
EBITDA_TA	0.06	0.05	0.13				
BLEV	0.29	0.18	0.32				
RET_TA	-0.12	0.00	0.42				
SIZE	14.14	14.10	1.47				
LIQUID	0.01	0.01	0.35				
EXPORT	0.23	0.00	0.42				
DIVPAY	0.28	0.00	0.45				
Panel B: Bankrupt firms							
EBITDA_TA	- 0.04	- 0.00	0.20				
BLEV	0.31	0.00	0.40				
RET_TA	- 0.76	- 0.28	1.06				
SIZE	13.59	13.60	1.92				
LIQUID	- 0.19	- 0.16	0.52				
EXPORT	0.15	0.00	0.35				
DIVPAY	0.09	0.00	0.28				
	EBIT DA_TA	BLEV	RET_TA	SIZE	LIQUID	EXPORT	DIV PAY

Panel C: *t* statistics testing for differences between means

Non-bankrupt versus bankrupt sample	60.76***	- 4.06***	107.48***	28.01***	41.54***	15.65***	32.64***
---	----------	-----------	-----------	----------	----------	----------	----------

This table presents the mean, median and standard deviation for the variables based on the sample of non-bankrupt firms (Panel A) and the sample of bankrupt firms (Panel B). EBITDA_TA is measured as earnings before interest, tax and depreciation divided by total assets. BLEV is measured as the book value of total debt to total assets. RET_TA is the ratio of retained earnings-to-total assets. SIZE is measured as the natural logarithm of total assets. LIQUID is defined as current assets minus current liabilities, all divided by total assets. EXPORT is a dummy variable that takes the value one if the firm exports, zero otherwise. DIVPAY is a dummy variable which equals to one if a firm pays out dividends, zero otherwise. Panel C reports *t* statistics testing the null hypothesis that there is no difference in the means between the non-bankrupt and bankrupt firms

***, ** and *Significance at the 1, 5 and 10% levels respectively

earnings-to-total assets and the probability of bankruptcy for US and UK SMEs respectively.

The column labelled BPM2 contains results from a model that augments the four firm-specific variables used as predictor variables in BPM1 with a liquidity ratio. The liquidity ratio significantly and negatively affects the probability of financial distress for private firms, suggesting that the more liquid a firm is the less likely it is to go bankrupt. This is consistent with the findings of Altman and Sabato (2007) for US SMEs and Altman et al. (2010) for UK SMEs. BPM3 reports the results from a model that augments BPM2 with a dummy variable that accounts for firms' export activity. The results show that export activity plays a significant role in the prediction of financial distress. The export dummy

Table 4 Results for multi-period logit models predicting the probability of financial distress

	BPM1	BPM2	BPM3	BPM4	BPM5
Panel A: Bankruptcy prediction models for Greek private firms					
Constant	- 3.8563*** (43.72)	- 3.7067*** (39.38)	- 3.9559*** (43.69)	- 3.8191*** (40.82)	- 3.2700*** (27.58)
EBITDA_TA	- 1.8623*** (25.42)	- 1.7627*** (22.86)	- 1.7480*** (22.70)	- 1.4407*** (18.20)	- 1.4071*** (14.64)
BLEV	0.6682*** (15.39)	0.5752*** (11.10)	0.5936*** (11.98)	0.5686*** (11.14)	0.5610*** (10.82)
RET_TA	- 0.8182*** (132.73)	- 0.7496*** (95.98)	- 0.7504*** (97.06)	- 0.7203*** (91.99)	- 0.7398*** (97.25)
SIZE	- 0.0949*** (5.02)	- 0.1039** (5.90)	- 0.0808** (4.11)	- 0.0796** (4.01)	- 0.0638** (4.25)
LIQUID		- 0.3563*** (6.99)	- 0.3559*** (7.01)	- 0.3589*** (7.29)	- 0.3409*** (7.18)
EXPORT			- 0.4364** (6.53)	- 0.3644** (4.54)	- 0.3658** (4.56)
DIVPAY				- 1.1838*** (20.66)	- 1.2220*** (21.91)
GRCRISIS					- 1.5655*** (22.75)
GDP GROWTH					- 0.1466*** (12.77)
Industry dummies	Yes	Yes	Yes	Yes	Yes
Log likelihood	- 8881.94	- 8860.47	- 8842.69	- 8758.43	- 8675.31
Pseudo R ² (%)	11.36	11.57	11.75	12.59	13.42
Wald statistic	2276.36***	2311.77***	2354.87***	2523.39***	2689.62***
Number of observations	188,065	188,065	188,065	188,065	188,065
Model <i>i</i> versus Model <i>j</i>					
Panel B: Likelihood ratio tests					
BPM5 versus BPM4					$\chi^2(2) = 27.27***$
BPM5 versus BPM3					$\chi^2(3) = 54.91***$

Table 4 continued

Model <i>i</i> versus Model <i>j</i>	<i>LR</i> statistic
BPM5 versus BPM2	$\chi^2(4) = 61.98^{***}$
BPM5 versus BPM1	$\chi^2(5) = 67.78^{***}$

This table contains estimation results for the multi-period logit models. 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. With the exception of the industry dummies, the variables are lagged to ensure that the data are observable prior to the event of financial distress. The column entitled BPM1 contains results for a model that uses size and three accounting ratios: profitability, leverage, and retained earnings divided by total assets. BPM2 contains the results from a model that incorporates the liquidity ratio along with the four variables in BPM1. BPM3 augments BPM2 with an export dummy variable while BPM4 contains results from a multi-period logit model that augments BPM3 with a dividend payout dummy variable. BPM5 contains results from a multi-period logit model that also incorporates a dummy variable for the Greek debt crisis and the real GDP growth rate. 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. 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). The row entitled Pseudo R^2 reports the pseudo R^2 for each model. Panel B contains results from likelihood ratio tests (scaled by the average number of observations per firm) examining whether model 5 contains incremental predictive information relative to model *j* for *j* = 1, ..., 4. The null hypothesis is that there is no difference between the two models which corresponds to testing the null hypothesis that the additional coefficients in model 5 relative to model *j* are jointly zero. The test follows a χ^2 distribution with *k* degrees of freedom, where *k* is the number of restrictions

***, ** and *Significance at the 1, 5 and 10% levels respectively

variable enters with a negative sign, which is as we expect given the earlier discussion in Sect. 2, and indicates that being involved in export activity reduces the probability of financial distress. The signs and the magnitude of the estimated coefficients on profitability, leverage, retained earnings-to-total assets, size and liquidity are similar to those from the model excluding the export dummy variable (BPM2). BPM4 augments BPM3 with a dividend payout dummy variable. The coefficient is statistically significant with a negative sign, indicating that a private firm that pays dividends is less likely to go bankrupt. This is perhaps not surprising as dividends can be used to convey information about a firm's future earnings and prospects. The inferences on the remaining five variables are qualitatively the same irrespective of the inclusion of the dividend payout dummy variable.

The final column headed BPM5 augments BPM4 with the real GDP growth rate and a dummy variable to capture the effects of the Greek debt crisis. We find that these two variables are significant predictors of the probability of financial distress. Intriguingly, we find that the crisis dummy variable is negatively related to the probability of financial distress. This means that bankruptcies of Greek private firms are less common during the crisis. The number of bankruptcies during the Greek debt crisis is lower than the number of bankruptcies prior to the Greek debt crisis. In particular, the number of bankruptcies in our sample is 88, 116 and 199 in 2009, 2010 and 2011 compared to 205, 348, 259, 157 and 398 for 2004 through 2008, for example. A possible explanation for this is that the number of firms that leave the dataset for other reasons (merger and acquisition, for example) increases during the crisis period in comparison to those firms that leave the dataset prior to the crisis period. Indeed, we observe in our sample that the number of firms that remain in the dataset decreases during the crisis period from 24,710 in 2008 to 21,984 in 2009, 20,546 in 2010 and 17,253 in 2011. The growth rate of real GDP also has a negative effect on the probability of financial distress, indicating that the probability of Greek private firms experiencing financial distress falls as real GDP rises.¹⁰ To summarise, then, the results in Table 4 indicate that there is some success to be had in using a multi-period logit model to identify significant predictors of the probability of private firms entering financial distress using financial statement information that has been found to predict the probability of financial distress for listed firms in developed and developing markets, as well as private firms in developed markets.

While the parameters are statistically significant in all of the models considered in Table 4, a natural question that arises is whether each model contains additional incremental predictive information. To investigate this, we compare the five bankruptcy prediction models to evaluate which model incorporates the most relevant information about the probability of financial distress. A comparison of each model's pseudo- R^2 shows that while BPM5 has the highest pseudo- R^2 of 14%, the lowest pseudo- R^2 , that for BPM1, is only a little lower at 11%. This raises the possibility that while the parameters in BPM5 are all statistically significant, one of the other more parsimonious models may perform just as well. Panel B of Table 4 reports likelihood ratio tests of the null hypothesis that model BPM5 contains no incremental predictive information over BPM1 through BPM4. Despite the fact that the pseudo- R^2 experiences only a reasonably modest increase from BPM1 to BPM5, the likelihood ratio tests strongly show that the additional variables in BPM5 relative to the other models have significant incremental predictive ability: in all cases, models 1 through 4 are rejected in favour of model 5.

¹⁰ Omitting the Greek crisis dummy from the model leads to a flip in sign for the real growth rate of GDP variable, which is counterintuitive.

Comparing our results for Greek private firms with the results for Greek listed firms documented in Charalambakis (2015) leads to some interesting conclusions. First, there are some common accounting ratios that are related to the probability of financial distress for both Greek listed and non-listed firms. In particular, leverage and profitability significantly predict the probability of financial distress, although the magnitudes of the coefficients for leverage and profitability are higher for Greek listed firms than for Greek private firms.¹¹ Second, we observe that variables that reflect a firm's financing constraints (firm size, liquidity, retained earnings-to-total assets and a firm's tendency to pay out dividends) strongly affect the probability of financial distress for Greek private firms. However, firm size and liquidity fail to predict financial distress for Greek listed firms. Third, inevitably, market data cannot be used for the prediction of the probability of financial distress for private firms. However, Charalambakis (2015) finds that past excess stock returns and stock return volatility are strongly associated with the probability of financial distress for Greek listed firms, a result consistent with findings for listed firms in developed markets (see, for example, Charalambakis and Garrett 2016 for the UK and Shumway 2001 for the US) though it contrasts with the findings in Charalambakis and Garrett (2016) for India, where Charalambakis and Garrett (2016) find that market-driven variables cannot predict financial distress for Indian listed firms.

As a final check on the performance of the models reported in Table 4 we follow Sobehart and Keenan (2001), Vassalou and Xing (2004) and Agarwal and Taffler (2008) among others and use the Area Under the Curve (AUC) criterion and the accuracy ratio (AR) to evaluate the models' fit in terms of predictive ability. The AUC is based on the area under the Receiver Operating Characteristics (ROC) curve and is an indicator of the quality of the model: the larger is the AUC, the better is the model at predicting the probability of financial distress. To arrive at an estimate of the AUC for each model, we first calculate the ROC curve. The ROC curve for each model BPM1 through BPM5 ranks firms from the highest estimated probability of default to the lowest and splits them into integers between 0 and 100 (these can be thought of as percentages). For each integer of the firms with the highest estimated probability of financial distress, we calculate the percentage of those firms that actually failed. These figures are then cumulated and the ROC curve plots them against the integer. If a model is incapable of distinguishing between bankrupt and non-bankrupt firms then the ROC curve will be a 45 degree line. The greater the predictive power of the model, the more bowed the curve will be and this is why the AUC forms the basis of a measure of the model's predictive ability and hence the quality of the model. If the bankruptcy prediction model is incapable of distinguishing between bankrupt and non-bankrupt firms (the ROC curve is a 45 degree line), the AUC is equal to 0.5; if the bankruptcy prediction model is perfectly capable of distinguishing between bankrupt and non-bankrupt firms the AUC is equal to 1. Engelmann et al. (2003) use the AUC to develop the accuracy ratio (AR) of the model which in turn forms the basis of a statistical test to evaluate whether the model performs better than one that, in this case, randomly allocates firms as being bankrupt or not. The accuracy ratio is defined as $AR = 2 \times (AUC - 0.50)$. An AR of zero shows that the model is incapable of performing better than a model which randomly allocates firms as being bankrupt or not. The higher the AR the better is the model at predicting the probability of bankruptcy. Results from the

¹¹ For Greek listed firms, profitability is marginally significant at the 10% level. In addition, Charalambakis (2015) finds that sales-to-total assets is strongly associated with the probability of financial distress for Greek listed firms but unreported results show that the ratio of sales-to-total assets is insignificant as a predictor of bankruptcy for Greek private firms.

Table 5 Area under the receiver operating characteristics (ROC) curve (AUC) and accuracy ratios

Models	AUC	SE	z	AR
BPM1	0.7984	0.0062	48.13	0.5968
BPM2	0.8061	0.0059	51.88	0.6122
BPM3	0.8042	0.0059	51.56	0.6084
BPM4	0.8115	0.0059	52.80	0.6230
BPM5	0.8191	0.0054	59.09	0.6382

This table presents the area under the ROC curve (AUC) for each of the multi-period logit models in Table 4. It also reports the standard error of the AUC (SE), z-statistics testing to see whether there is a significant difference between the AUC for the relevant multi-period logit model and that for a model that randomly allocates firms as being bankrupt or not (this latter model has an AUC of 0.50), and the accuracy ratio (AR) of each model. The row entitled BPM1 contains results for a model that uses size and three accounting ratios: profitability, leverage, and retained earnings divided by total assets. BPM2 presents the results from a model that incorporates the liquidity ratio along with the four variables in BPM1. BPM3 augments BPM2 with an export dummy variable while BPM4 contains results from a multi-period logit model that augments BPM3 with a dividend payout dummy variable. BPM5 contains results from a multi-period logit model that also incorporates a dummy variable for the Greek debt crisis and the real GDP growth rate

estimated AUC and the corresponding ARs are presented in Table 5. We also report a z-statistic, defined as $z = \frac{\widehat{AUC} - 0.5}{\text{std.error}(AUC)}$, testing the null hypothesis that there is no difference in the performance of the bankruptcy prediction model and a random classification model with AUC equal to 0.5. The z statistics reported in Table 5 are all large and statistically significant, showing that all five models outperform the random classification model. This is also confirmed by the AR as in all cases the AR is greater than zero. However, the model with the largest AUC and the highest accuracy ratio is BPM5, a finding that is consistent with the results in Table 4.

4.2 In-sample and out-of-sample forecast accuracy

The results in the previous subsection suggest that it is possible to predict the probability of financial distress for Greek private firms using accounting ratios and that the models perform significantly better than one which randomly allocates firms as bankrupt or not. What we have yet to see, however, is how well the models correctly classify firms that do go bankrupt as likely to go bankrupt. In order to evaluate this, we sort firms in descending order based on the probability of bankruptcy estimated by each of the five multi-period logit models described and discussed in the previous subsection. Deciles 1 through 5 contain firms that are predicted as more likely to enter financial distress, while deciles 6 through 10 contain those firms that are less likely to enter financial distress. We define the percentage of bankrupt firms that are allocated to the various groups by the estimated probability of financial distress derived from each model. This is a means by which we can assess the ability of the models to correctly classify those firms that went bankrupt as likely to go bankrupt. In particular, for each model, we report the percentage of bankrupt firms that are classified as firms with a higher probability of financial distress (these firms should be placed in deciles 1 through 5) and the percentage of bankrupt firms classified as firms with a lower probability of financial distress (deciles 6 through 10). These represent the

Table 6 In-sample forecast accuracy tests

Decile	BPM1	BPM2	BPM3	BPM4	BPM5
1	52.88	53.56	53.22	53.38	54.52
2	15.14	14.92	15.03	15.37	15.14
3	7.74	8.47	7.74	7.63	7.80
4	5.54	5.31	5.48	6.27	5.48
5	3.84	4.12	4.01	3.96	4.12
6–10	14.86	13.62	14.52	13.39	12.94
No. of bankrupt firms	1770	1770	1770	1770	1770

This table examines the forecast accuracy of the five multi-period logit models we estimate and whose results are reported in Table 4. Firms are sorted into deciles based on their estimated probability of financial distress. Deciles 1–5 contain those firms with the highest probability of financial distress while Deciles 6–10 contain those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt that the model places into each decile. BPM1 contains results for a model that uses size and three accounting ratios: profitability, leverage, and retained earnings divided by total assets. BPM2 presents the results from a model that incorporates the liquidity ratio along with the four variables in BPM1. BPM3 augments BPM2 with an export dummy variable while BPM4 contains results from a multi-period logit model that augments BPM3 with a dividend payout dummy variable. BPM5 contains results from a multi-period logit model that also incorporates a dummy variable for the Greek debt crisis and the real GDP growth rate

classification and misclassification rates respectively of each model. Table 6 reports the ability of the models to correctly classify firms that went bankrupt as being likely to go bankrupt when the models are estimated using the entire sample. These results can be thought of as in-sample forecast tests. The results in Table 6 clearly show that all of the models perform well in terms of correct classification and misclassification, with all of the models correctly allocating 85% or more of firms that actually went bankrupt in to deciles 1 through 5. The model that exhibits the best predictive ability in terms of correctly classifying firms as bankrupt is BPM5, classifying 87.06% of firms that went bankrupt into deciles 1 through 5. However, while these in-sample results lend strong support to the models in Table 4, and BPM5 in particular, as predictors of the probability of financial distress there is a danger that the models are uninformative when forecasting out-of-sample due to over-fitting, a situation in which a model includes predictors which improve the in-sample fit of the model but penalize the model when forecasting out-of-sample.

To investigate whether this is the case, we examine the ability of the models in Table 4 to predict the probability of financial distress out-of-sample as this should be the ultimate criterion when evaluating bankruptcy prediction models. We re-estimate the multi-period logit models described both above and in Table 4 using data from 2003 to 2008, a period prior to the Greek debt crisis. Since the estimation sample stops prior to the Greek debt crisis and we are forecasting out-of-sample, we do not include the dummy variable for the Greek debt crisis in BPM5.¹² We then use the estimated coefficients from each model to predict corporate bankruptcies throughout the Greek debt crisis period of 2009–2011. The results are shown in Table 7. We observe that our preferred model, BPM5, exhibits the best out-of-sample performance. It classifies 88.34% of firms that went bankrupt during the period of the financial crisis into deciles 1 through 5 and although it has the same

¹² Recall from earlier that the Greek debt crisis dummy variable takes on the value 1 from 2009 to 2011, zero otherwise.

Table 7 Out-of-sample forecast accuracy tests

Decile	BPM1	BPM2	BPM3	BPM4	BPM5
1	55.09	56.08	57.57	57.83	58.07
2	15.38	15.38	11.66	11.41	10.92
3	7.94	6.20	7.69	7.94	8.19
4	4.22	5.46	5.21	6.20	6.20
5	5.21	4.72	4.96	4.96	4.96
6–10	12.16	12.16	12.91	11.66	11.66
No. of bankrupt firms	403	403	403	403	403

This table examines the out-of-sample forecast accuracy of the multi-period logit models we estimate. The models are estimated using data over the period 2003–2008. These parameter estimates are then used to calculate the probability of financial distress over the period 2009–2011. Firms are sorted into deciles based on their estimated probability of financial distress. Deciles 1–5 contain those firms with the highest probability of financial distress while Deciles 6–10 contain those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt that the model places into each decile. BPM1 contains results for a model that uses size and three accounting ratios: profitability, leverage, and retained earnings divided by total assets. BPM2 presents the results from a model that incorporates the liquidity ratio along with the four variables in BPM1. BPM3 augments BPM2 with an export dummy variable while BPM4 contains results from a multi-period logit model that augments BPM3 with a dividend payout dummy variable. BPM5 contains results from a multi-period logit model that also incorporates the real GDP growth rate. The dummy variable for the Greek debt crisis that takes on the value 1 for the period 2009–2011, zero otherwise, is omitted

misclassification rate as BPM4, BPM5 correctly classifies a slightly higher percentage of bankrupt firms into the first decile. These results suggest that our best model in terms of in-sample predictive performance, BPM5, works well in forecasting bankruptcy not only out-of-sample but also during the crisis period. However, the possibility remains that while the model works well over the whole crisis period, it may not work so well for each year of the crisis. To investigate whether the out-of-sample predictive ability of BPM5 sustains for each year of the crisis,¹³ we estimate BPM5 with data from 2003 to 2008 and use the estimated coefficients to predict bankruptcy in 2009. We then re-estimate the model extending the sample until 2009 and use the model to predict the probability of financial distress in 2010. Finally, we re-estimate the model using data until 2010 and use the estimated parameters to predict the probability of financial distress in 2011. For each model, we report the percentage of firms that went bankrupt that are classified by the model as firms that have a higher probability of financial distress, that is, we report the percentage of those firms that went bankrupt that the model places in deciles 1 through 5, for each corresponding year of the crisis. We find that the model correctly classifies 87.5% of bankrupt firms in 2009, 89.66% of bankrupt firms in 2010 and 87.94% of bankrupt firms in 2011. Overall, the out-of-sample predictive ability of our model is high for both the crisis period as a whole and each year within the crisis period.

4.3 Extending the forecast horizon

In this subsection we explore the ability of our best model, BPM5, to predict the probability of financial distress when we extend the forecast horizon from one year to two and

¹³ We would like to thank an anonymous referee for suggesting this.

Table 8 Extending the forecast horizon

	2-year forecast horizon	3-year forecast horizon
Constant	- 4.9595*** (49.86)	- 4.5039*** (27.67)
EBITDA_TA	- 1.3925*** (9.97)	- 1.0796** (5.01)
BLEV	0.5397** (7.74)	0.5138** (5.64)
RET_TA	- 0.7138*** (64.65)	- 0.7475*** (48.57)
SIZE	- 0.0249 (1.51)	- 0.0281 (1.30)
LIQUID	- 0.3207** (4.94)	- 0.2991** (4.76)
EXPORT	- 0.3382** (4.75)	- 0.2775** (4.35)
DIVPAY	- 0.9961*** (16.80)	- 0.9110*** (10.99)
GRCRISIS	- 0.5667*** (6.78)	- 0.9862*** (27.06)
GDP GROWTH	0.0144 (0.92)	- 0.2117*** (19.31)
Industry dummies	Yes	Yes
Log likelihood	- 7256.25	- 5180.47
Wald statistic	1632.15***	1185.83***
Pseudo R^2	0.1011	0.1027
Number of observations	157,556	128,378

This table presents the results of estimating BPM5 when the forecast horizon is increased from one to two and three years, respectively. 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. 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). The row entitled Pseudo R^2 reports the pseudo R^2 for each model

***, ** and *Significance at the 1, 5 and 10% levels respectively

three years. In order to do this, we use (with the exception of the dummy variables) \mathbf{x}_{it-2} and \mathbf{x}_{it-3} rather than \mathbf{x}_{it-1} in (2). The results are presented in Table 8. As can be seen, the signs and the magnitudes of the estimated coefficients on the variables are similar to those documented in Table 4. Profitability, retained earnings-to-total assets, the export dummy variable, the dividend payout dummy variable, the crisis dummy variable and the GDP growth rate are negatively related to the probability of financial distress while leverage is positively related to the probability of financial distress. The only notable difference is that when we extend the forecast horizon, size is no longer significant, suggesting that size is important for predicting the probability of financial distress over the short term (one year) but it's importance subsides over longer forecasting horizons. Therefore, with this one exception, the variables that were found to predict the probability of financial distress in Table 4 remain strongly associated with the probability of financial distress over longer forecast horizons.

Table 9 reports results from repeating the in-sample and out-of-sample analysis of Tables 6 and 7 for the 2 and 3-year forecast horizons. The interesting finding here is that the forecast accuracy of the model remains strong. As expected, the forecast accuracy of the model decreases when we increase the forecast horizon. However, BPM5 still performs well. Looking at Panel A of Table 9, we observe that the misclassification rate of the model when forecasting out-of-sample increases from 11.66% when forecasting 1-year-ahead to 13.75 and 16.55% for a 2 and a 3-year forecast horizon respectively. This means that the model correctly classifies 86.25% of bankrupt firms for a 2-year horizon and 83.45% of

Table 9 Out-of-sample and in-sample forecast accuracy tests when extending the forecast horizon

Decile	2-year horizon	3-year horizon
Panel A: Out-of-sample forecast tests		
1	53.87	48.65
2	11.46	11.49
3	7.16	9.12
4	6.88	8.11
5	6.88	6.08
6–10	13.75	16.55
No. of bankrupt firms	349	296
Panel B: In-sample forecast tests		
1	46.74	43.60
2	14.43	14.84
3	10.18	8.94
4	6.51	7.93
5	5.66	5.59
6–10	16.48	19.10
No. of bankrupt firms	1414	984

Panel A presents the out-of-sample forecast results from the best multi-period logit model (BPM5) when the forecast horizon is two and three years, respectively. For the out-of sample forecast tests the model is estimated using data over the period 2003–2008 (the Greek debt crisis dummy variable is omitted). The parameter estimates are then used to calculate the probability of financial distress over the period 2009–2011. For the in-sample tests the model is estimated using the full sample and the Greek debt crisis dummy is included in the model. Firms are then sorted into deciles based on their estimated probability of financial distress. Deciles 1–5 contain those firms with the highest probability while Deciles 6–10 contain those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt that the model places in to each decile. Panel B presents the in-sample performance of the multi-period logit model when the forecast horizon is two and three years, respectively. Firms are sorted into deciles based on their estimated probability of financial distress. Deciles 1–5 contains those firms with the highest probability while Deciles 6–10 contain those with the lowest. We then calculate the percentage (to two decimal places) of firms that subsequently went bankrupt that the model places in to each decile

bankrupt firms into deciles 1 through 5. For completeness the in-sample results are presented in Panel B of Table 9. The in-sample test documents that the model allocates more than 80% of bankrupt firms to deciles 1 through 5 at both the 2 and 3-year horizons (83.52 and 80.9% respectively).

4.4 Additional robustness checks

The first robustness check we perform is related to the in-sample and out-of-sample forecasting ability of our preferred model discussed in Sect. 4.2. In terms of ability to classify bankrupt firms as having higher probabilities of going bankrupt, BPM5 works well both in-sample and out-of-sample and seems to do a decent job of predicting the probability of financial distress over the Greek debt crisis period. However, while our earlier results show that BPM5 forecasts well, a question we did not directly address is whether the nature of the relationship between the predictor variables and the probability of financial distress is substantively different once we allow for the effects of the Greek debt

Table 10 Bankruptcy prediction models including and excluding the Greek crisis period

	BPM5, Full sample	BPM5 Excluding the Greek crisis
Constant	− 3.2700*** (27.58)	− 3.7856*** (41.87)
EBITDA_TA	− 1.4071*** (14.64)	− 1.5552*** (20.32)
BLEV	0.5610*** (10.82)	0.6509*** (17.00)
RET_TA	− 0.7398*** (97.25)	− 0.7429*** (111.59)
SIZE	− 0.0638** (4.25)	− 0.0372 (0.82)
LIQUID	− 0.3409*** (7.18)	− 0.2690** (4.68)
EXPORT	− 0.3658** (4.56)	− 0.3360* (3.82)
DIVPAY	− 1.2220*** (21.91)	− 1.2235*** (28.65)
GRCRISIS	− 1.5655*** (22.75)	
GDP GROWTH	− 0.1466*** (12.77)	− 0.1142*** (7.48)
Industry dummies	Yes	Yes
Log likelihood	− 8675.31	− 66618.15
Pseudo R^2 (%)	13.42	12.55
Wald statistic	2689.62***	1899.84***
Number of observations	188,065	126,282

The second column reproduces the results for BPM5 from Table 4 while the third column presents the results from BPM5 when the sample period runs from 2003 to 2008. 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. 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). The row entitled Pseudo R^2 reports the pseudo R^2 for each model

***, ** and *Significance at the 1, 5 and 10% levels respectively

crisis. In other words, to what extent are the results in Table 4 driven by the Greek debt crisis? Table 10 reports results from estimating the model using the full sample and from estimating the model using a sample running from 2003 to 2008. With respect to the pre-Greek-crisis period, the estimated parameters have the same sign and are of similar magnitudes to those from estimating the model using the entire sample period with one exception: that of size. In the pre-Greek-debt-crisis sample period, the coefficient on size becomes much smaller and statistically insignificant. This suggests that outside of a crisis, size is relatively unimportant in predicting the probability of financial distress relative to the other variables.

A second robustness check we perform is to determine how the model performs when we split firms into small-sized firms and medium-sized firms. Although our earlier models contain the log of total assets as a proxy for size, this is not how the European Union defines size. Therefore, following the recommendation of the Commission (2003), we define the size of a firm based on the number of employees. A small firm is considered to be a firm with less than fifty employees while a medium-sized firm is a firm with fifty to 249 employees. This leads to samples containing 26,534 small firms with 163,851 firm-year observations and 2460 medium-sized firms with 17,315 firm-year observations. Table 11 reports the results from estimating BPM5 using these two samples. A notable difference in the results for small firms compared to the results for the whole sample is the insignificance of the export dummy variable for small firms: there is no

Table 11 Bankruptcy prediction when small- and medium-sized firms are separated

	Small firms	Medium firms
Constant	− 6.6248*** (47.13)	− 10.9983*** (6.92)
EBITDA_TA	− 1.0507** (4.29)	− 1.9586 (2.56)
BLEV	0.5479*** (5.85)	1.4435** (4.00)
RET_TA	− 0.7358*** (46.17)	− 0.8853** (5.67)
SIZE	− 0.0118 (0.19)	0.2218 (1.16)
LIQUID	− 0.5344*** (7.83)	− 0.4010 (1.34)
EXPORT	0.0496 (0.30)	− 0.6689** (4.45)
DIVPAY	− 1.3641*** (15.26)	− 0.9109* (2.84)
GRCRISIS	− 2.3919*** (25.79)	− 2.9695** (3.90)
GDP GROWTH	− 0.1909*** (11.18)	− 0.2781** (3.88)
Industry Dummies	Yes	Yes
Log Likelihood	− 5228.96	− 443.54
Wald statistic	1352.81***	143.29***
Pseudo R^2	0.1145	0.1391
Number of observations	163,613	17,276

This table presents the results from estimating the best multi-period logit model (BPM5) for small and medium-sized firms separately. Following the recommendation of the Commission (2003), we define the size of a firm based on the number of employees. A firm is considered to be small if it has less than 50 employees while a medium-sized firm is a firm with 50–249 employees. 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. 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). The row entitled Pseudo R^2 reports the pseudo R^2 for each model

***, ** and *Significance at the 1, 5 and 10% levels respectively

statistically significant association between export activity and the probability of financial distress for small firms. For medium-sized firms, leverage, retained earnings-to-total assets, the export dummy variable and the dividend payout dummy variable are significant and have the expected signs. However, profitability and liquidity are not significant predictors of financial distress for medium-sized firms. The results also reveal that the predictive power of leverage and the GDP growth rate increases for medium-sized firms compared to small firms while the predictive power of the liquidity ratio and dividend payout is stronger for small firms than medium firms.

A final robustness test we consider is whether the inclusion of macroeconomic variables beyond the growth rate in real GDP affects the probability of financial distress for private firms. We augment BPM5 with three macroeconomic variables: the term spread, defined as the difference between the 10-year Greek government bond yield and the 10-year German government bond yield, domestic credit to the private sector scaled by GDP, and Greek public debt scaled by GDP. None of these additional macroeconomic variables are statistically significant.

5 Concluding remarks

In spite of their economic importance, little is known about what determines the probability of financial distress for private firms, particularly in a developing economy. This paper examines the probability of financial distress for private firms in a developing economy using the discrete hazard/multi-period logit approach (Shumway 2001). Using a large dataset of Greek private firms covering the period from 2003 to 2011 we find that seven firm-specific factors, many of which have been found to successfully predict the probability of financial distress for public firms in both developed and developing economies, as well as for SMEs in the US and UK, are strongly associated with the probability of financial distress for private firms in Greece. Profitability, retained earnings-to-total assets, size, liquidity, a dummy variable for export activity and a dividend payout dummy variable are negatively associated with the probability of bankruptcy whereas leverage is positively related to the probability of bankruptcy. We further find that after controlling for the Greek debt crisis, the real GDP growth rate has a significantly negative impact on the probability of financial distress and therefore matters when forecasting financial distress for Greek private firms. A model that controls for the Greek debt crisis and incorporates these variables exhibits the highest predictive ability based on both in-sample and out-of-sample forecast accuracy tests. The model also retains its predictive power when we increase the forecast horizon from one to two and three years. To examine the extent to which our findings are driven by the Greek debt crisis, we re-estimated our preferred model excluding the Greek debt crisis. With the exception of size, which becomes insignificant, the estimated parameters have the same sign and are of similar magnitude to the full-sample estimates. The model also correctly classifies 88% of firms that went bankrupt during the Greek debt crisis as likely to go bankrupt. We also find that when we split firms in to size according to the number of employees, the impact of the variables on the probability of financial distress can vary across small and medium-sized firms.

Acknowledgements We would like to thank the anonymous referees for helpful and thoughtful comments that have improved the paper. We would also like to thank Heather Gibson and seminar participants at the Economic Research Department of the Bank of Greece for their valuable comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agarwal V, Taffler RJ (2008) Comparing the performance of market-based and accounting-based bankruptcy prediction models. *J Bank Finance* 32:1541–1551
- Alifiah MN (2014) Prediction of financial distress companies in the trading and services sector in Malaysia using macroeconomic variables. *Proc Soc Behav Sci* 129:90–98
- Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J Finance* 23:589–609
- Altman EI, Haldeman RG, Narayanan P (1977) Zeta analysis: a new model to identify bankruptcy risk of corporations. *J Bank Finance* 1:29–54
- Altman EI, Sabato G (2007) Modelling credit risk for SMEs: evidence from the US market. *Abacus* 43:332–357

- Altman EI, Sabato G, Wilson N (2010) The value of non-financial information in small and medium-sized enterprise risk management. *J Credit Risk* 6:1–33
- Asker J, Farre-Mensa J, Ljungqvist A (2012) Comparing the investment behavior of public and private firms. NBER working paper no. 17394
- Bauer J, Agarwal V (2014) Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *J Bank Finance* 40:432–442
- Beaver WH (1966) Financial ratios as predictors of failure. *J Acc Res* 4:71–111
- Bernard AB, Jensen JB (1995) Exporters, jobs, and wages in US manufacturing: 1976–1987. In: Baily MN, Reiss PC, Winston C (eds) *Brookings papers on economic activity, Microeconomics*. Brookings Institution Press
- Bernard AB, Jensen JB (1999) Exceptional exporter performance: cause, effect, or both? *J Int Econ* 47:1–25
- Bernard AB, Jensen JB, Redding SJ, Schott PK (2007) Firms in international trade. *J Econ Perspect* 21:105–130
- Bharath ST, Shumway T (2008) Forecasting default with the Merton distance to default model. *Rev Financ Stud* 21:1339–1369
- Bhattacharya S (1979) Imperfect information, dividend policy, and “The Bird in the Hand” fallacy. *Bell J Econ* 10:259–270
- Brav O (2009) Access to capital, capital structure and the funding of the firm. *J Finance* 64:263–308
- Campbell JY, Hilscher T, Szilagyi J (2008) In search of distress risk. *J Finance* 63:2899–2939
- Cangemi B, Servigny De A, Friedman C (2003) Standard and poor’s credit risk tracker for private firms. Technical Document, Standard & Poor’s Risk Solutions
- Charalambakis EC (2015) On the prediction of corporate financial distress in the light of the financial crisis: empirical evidence from Greek listed firms. *Int J Econ Bus* 22:407–428
- Charalambakis EC, Garrett I (2016) On the prediction of financial distress in developed and emerging markets: does the choice of accounting and market information matter? A comparison of UK and Indian firms. *Rev Quant Financ Acc* 47:1–28
- Chava S, Jarrow RA (2004) Bankruptcy prediction with industry effects. *Rev Finance* 8:537–569
- Diekes M, Erner C, Langer T, Norden L (2013) Business credit information sharing and default risk of private firms. *J Bank Finance* 37:2867–2878
- Duffie D, Saita L, Wang K (2007) Multi-period corporate failure prediction with stochastic covariates. *J Financ Econ* 83:635–665
- Engelmann B, Hayden E, Tasche D (2003) Testing rating accuracy. *Risk* 16:82–86
- Etheridge HL, Sriram RS (1997) A comparison of the relative costs of financial distress models: artificial neural networks, logit and multivariate discriminant analysis. *Intell Syst Acc Finance Manag* 6:235–248
- Commission European (2003) Commission recommendation 2003/361/EC concerning the definition of micro, small and medium-sized enterprises. *Off J Eur Union* 46:36–41
- Faccio M, Marchica M-T, McConnell J, Mura R (2012) Returns and risks to private equity. Available at SSRN: <https://doi.org/10.2139/ssrn.1781070>
- Falkenstein E, Boral A, Carty L V (2000) RiskCalc for private companies: Moody’s default model. As published in *Global Credit Research*, May 2000. Available at SSRN: <http://ssrn.com/abstract=236011> or <https://doi.org/10.2139/ssrn.236011>
- Gao H, Hartford J, Li K (2012) CEO pay cuts and forced turnover: their causes and consequences. *J Corp Finance* 18:291–310
- Garrett I, Priestley R (2000) Dividend behavior and dividend signalling. *J Financ Quant Anal* 35:173–189
- Hillegeist SA, Keating EK, Cram DP, Lundstedt KG (2004) Assessing the probability of bankruptcy. *Rev Acc Stud* 9:53–85
- Kwak W, Shi Y, Kou G (2012) Bankruptcy prediction for Korean firms after the 1997 financial crisis: using a multiple criteria linear programming data mining approach. *Rev Quant Financ Acc* 38:441–453
- Lau AH (1987) A five-state financial distress prediction model. *J Acc Res* 25:127–138
- Lintner J (1956) Distribution of incomes of corporations among dividends. *Am Econ Rev* 61:97–113
- Marsh T, Merton RC (1987) Dividend behavior for the aggregate stock market. *J Bus* 60:1–40
- Merton RC (1974) On the pricing of corporate debt: the risk structure of interest rates. *J Finance* 29:449–470
- Michaela R, Roberts MR (2012) Corporate dividend policies: lessons from private firms. *Rev Financ Stud* 25:711–746
- Miller MH, Rock K (1985) Dividend policy under asymmetric information. *J Finance* 40:1031–1051
- Nittayagasetwat A (1996) A use of untransformed financial accounts in a neural network model for predicting bankruptcy. *NIDA Dev J* 36:81–108
- OECD (2015) *Financing SMEs and entrepreneurs 2015: an OECD scoreboard*. OECD Publishing, Paris
- Ohlson J (1980) Financial ratios and the probabilistic prediction of bankruptcy. *J Acc Res* 18:109–131

- Petersen MA, Rajan RG (1994) The benefits of lending relationships: evidence from small business data. *J Finance* 49:3–37
- Sobehart J, Keenan S (2001) Measuring default accurately. *Risk* 14:31–33
- Shumway T (2001) Forecasting bankruptcy more accurately: a simple hazard model. *J Bus* 74:101–124
- Taffler RJ (1983) The assessment of company solvency and performance using a statistical model. *Acc Bus Res* 13:295–307
- Tinoco MH, Wilson N (2013) Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *Int Rev Financ Anal* 30:394–419
- Vassalou M, Xing Y (2004) Default risk in equity returns. *J Finance* 59:831–868
- Van Biesebroeck J (2005) Exporting raises productivity in sub-Saharan African manufacturing firms. *J Int Econ* 67:373–91
- Wymenga P, Spanikova V, Barker A, Konings J, Canton E (2012) EU SMEs in 2012: at the crossroads. Annual report on small and medium-sized enterprises in the EU, 2011/12, Rotterdam
- Zmijewski M (1984) Methodological issues related to the estimation of financial distress prediction models. *J Acc Res* 22:59–82