Estimating the Probability of Financial Distress:

International Evidence

Julio Pindado*
Universidad de Salamanca, Spain
Luis Rodrigues
Escola Superior de Tecnologia de Viseu, Portugal
Chabela de la Torre
Universidad de Salamanca, Spain

Abstract

This study focuses on developing a new approach to estimating the ex-ante probability of financial distress by means of a model that could be applied to different economic and legal contexts. Our approach first consists of testing for the specification of the proposed model by using panel data methodology to eliminate the unobservable heterogeneity. Second**,** the model is cross-sectionally estimated to obtain an indicator of the probability of financial distress that incorporates the specificity of each company. We find that such a probability is accurately explained by a smaller number of theoretically underpinned factors than has been generally assumed.

EFM *classification*: 130 Bankruptcy and Financial Distress

Keywords: Financial insolvency, probability of financial distress, logit analysis

1. INTRODUCTION

Since the seminal article of Altman (1968), scholars have developed many models for predicting financial distress and/or bankruptcy, which have been widely applied as evaluation models providing credit risk information since the mid-1970s. Altman's first model (Z-score) has undergone several revisions, and many alternative models using different variables and techniques have been suggested. The research by Taffler (1983, 1984) for the UK is a good example of the development and application of the seminal model to different countries. Altman (1984b), Jones (1987), Altman and Narayanan (1997) and Altman and Saunders (1997) offer several surveys describing other model designs, innovations and outcomes in this strand of literature. For example, Altman et al. (1977) proposed the Zeta model and, in the 1980s, the Ohlson's (1980) proposal to substitute Logistic Analysis (LA) for Linear Discriminant Analysis (LDA) as the estimation method was followed by several authors, such as Zavgren (1985). Following the same trend Zmijewski (1984) opted for a Probit Analysis[1].

The international financial crisis that began in Asia in 1997 also encouraged renewed research on this topic. Recent studies carried out inside international financial institutions (The World Bank and IMF), such as those by Claessens and Klapper (2003) and Claessens et al. (2003), show an increasing concern with companies' financial distress that can lead to a domino effect, triggering financial crisis at an international level.

In this context, our aim is to develop a new approach to estimating a probability of financial distress that can be incorporated into credit scoring models, reducing their instability over time. As already discussed by Barnes (1987) and Ward and Foster (1997), who criticized the dominance of legal bankruptcy in financial distress research, there has been a lack of consensus among scholars on a criterion of distress. Hence our study focuses on financial distress regardless of the legal consequences of this situation,

[1] A comparison of the mentioned models concerning methodologies, number of companies, period of estimation, and type of variables can be found in Appendix 1.

because bankruptcy is a legally, rather than an economically, defined event. Therefore, we opt for a financial criterion related to a situation of not being able to comply with financial obligations when they become due, which is assumed to be attained when the earnings before interests, taxes and amortization ($EBITDA_{it}$) are smaller than financial expenses.

Additionally, this criterion of distress is consistent with our ex-ante approach. That is, we have chosen an ex-ante approach in order to overcome the criticism made by Wood and Piece (1987) of the use of ex-post models when predicting financial distress ex-ante, since we are interested in estimating the probability of financial distress faced by all firms regardless of whether they file for bankruptcy or not. Note that the ex-ante distress costs are very important in that they are born by all firms (White, 1996) and, consequently, these costs influence a firm's financial decisions. Therefore, this approach to estimating the probability of financial distress faced by each firm is a very important instrument for both firms and research scholars. Essentially, our purpose is to build an indicator of the probability of financial distress that can advantageously be used not only in credit-scoring applications, but also to test the role played by financial distress costs in economic theory.

To achieve our aim, our paper offers a new empirical approach leading to a better judgment of the PFD. The selection of explanatory variables has been traditionally based on sequential processes of elimination of variables according to a maximum prediction capacity criterion. This empirical method often leads to over-adjusted models with counter-intuitive coefficient signs and results. In contrast, our basic idea essentially consists of developing a financial distress probability model according to financial theory, and testing for its specification by using panel data methodology. This methodology allows us to eliminate the unobservable heterogeneity, and to solve the problem of choosing the estimation year before the crisis by using the maximum annual data for each firm and thus improving the accuracy of the model. Once the correct specification of our model has been tested, it is used to perform a cross-sectional analysis to predict a PFD that incorporates the specificity of each company.

Our empirical evidence validates the econometric specification of the proposed model. The estimated coefficients in panel data models yield the expected sign for all countries using both fixed and random effect panel data methodologies. Specifically, we find that the PFD is significantly explained by a number of theoretically underpinned factors that is smaller than has been generally assumed, namely the company's returns on assets, and the trade-off between this way of generating funds and the company's need to comply with its financial expenses during the financial year. The results of the cross-sectional analysis strongly confirm the accuracy of our model, and show percentages of correct classification above 90 per cent for all countries and years.

Our approach to estimating the PFD can be applied in order to improve other models that have been developed in several areas of finance, such as the effect of financial distress costs on financial structure (Mackie-Mason, 1990; Graham, 1996; Leary and Roberts, 2004), by using a method that does not rely on an automatic classification of the level of financial distress according to debt as occurs in Opler and Titman (1994).

The remainder of the paper is organized as follows. Section 2 describes the data and variables used in our study. In Section 3, a parsimonious model for estimating the PFD is specified, for which we then propose an innovative strategy of estimation that incorporates panel data methodologies as well as cross-sectional analyses in Section 4. Throughout Section 5 we present and comment the estimation results of our PFD model. Finally, Section 6 concludes the paper.

2. DATA: SAMPLES AND VARIABLES

According to the new approach we propose here, data from a group of developed countries are needed in order to make sure that our model for the PFD works regardless of the data used to estimate it. We have thus used an international database, the Compustat Global Vantage, as our source of information.

For each country we constructed a panel of firms with information for at least six consecutive years during the period from 1990 to 1999. There are only a few countries for which samples with the mentioned structure can be selected. Specifically, we have been able to select samples with enough size to test our hypothesis for Germany, the United Kingdom (UK) and the United States (US). Note that these countries represent a great variety of institutional environments. The distribution by number of companies and number of annual observations per country is provided in Table 1.

We have thus constructed an unbalanced panel with between six and ten year of data for each company, combining the available Compustat Global Industrial Active files containing information on active companies, with Compustat Global Industrial Research files which provide data on companies which were suspended from quotation for some reason after a certain period in the capital market[2]. This data structure allows the number of observations to vary across companies and thus represents added information for our model. This way we can use the largest number of observations, which reduces the possible survival bias arising when the observations in the initial cross-section are independently distributed and subsequent entries and exits in the panel

[2] Firms that filed for bankruptcy are an example. However, companies in such a situation only represent a small percentage of the available data and, even in these cases, the available information is of poor quality as a natural consequence of the degradation of the flow of information characterizing severe crises.

occur randomly. The financial companies in Compustat Global (SIC code 5000) were excluded companies because they have their own specificity.

Since the available information in Compustat Global on companies that filed for bankruptcy is rather scarce for the type of study we intend to put forward, we propose a concept of financial distress that, in essence, is similar to a “situation of a firm which can no longer meet its financial obligations as they become due” (Beaver, 1966). Our view is that a company reaches this critical point when its EBITDA is smaller than its financial expenses. Previous research widely supports this definition of financial distress, such as Wruck (1990), Asquith et al. (1994), Andrade and Kaplan (1998) and Whitaker (1999).

This concept can be universally applied, as opposed to bankruptcy which is legal and country specific[3], and it allows us to use all the companies presented in our country panels in the estimation of the PFD model. In this way, we avoid the traditional criticism of the use of samples of reduced size defined by a rather strict concept of financial distress.

Finally, we have selected a smaller set of variables explaining the PFD, whose theoretical foundation is explained in Section 3. First, Earnings Before Interests and Taxes ($EBIT_{it}$ /K_{it-1}), which contains information about how a company is able to efficiently use its assets in order to generate the necessary funds during the financial year. Second, Financial Expenses (FE_{it}/K_{it-1}) has been chosen instead of debt stock ratios because the latter seem to lose explanatory power as compared to the chosen flow variable. Third, we also adopt a variable of historical profitability (RE_{it}/K_{it-1}), since it has been recognized as one of the most accurate variables in financial crisis detection.

[3] Additionally, it is worthwhile noting that this kind of definition based on bankruptcy cannot be applied when predicting financial distress ex-ante.

Moreover, all these variables are scaled according to the replacement value of total assets at the beginning of the period (K_{it-1}) instead of the book value of total assets.

3. A MODEL FOR THE PROBABILITY OF FINANCIAL DISTRESS

Our goal is to estimate a logistic model that can provide us with an indicator of the PFD. We have already mentioned the importance of building a study upon a concept of financial distress that does not rely on legal institutions, but follows Beaver's (1966) intuition.

Another noteworthy feature of our study refers to the selection of the independent variables in the model. According to Scott (1981), the selection of explanatory variables should not be based on sequential processes of elimination of variables according to a maximum prediction capacity criterion, since this method often leads to over-adjusted models with counter-intuitive coefficient signs and results. Consequently, we have selected our explanatory variables relying on a strong theoretical justification. Note that these variables showed the highest discriminatory power in the models of Altman (1968), Altman et al. (1977), and Ohlson (1980), as well as in the subsequent studies by Begley et al. (1996), Dichev (1998) and Cleary (1999). Additionally, this selection of variables allows us to specify a logistic model that is intended to be stable and parsimonious and that reduces the problems concerning the choice of economic contexts and periods to consider.

3.1. A financial-based definition of financial distress

In contrast to most of the literature, our study is focused on financial distress regardless of the legal consequences of this situation. We focus on financial distress instead of bankruptcy prediction, since, as Barnes (1987, 1990) points out, the failure to meet financial obligations does not necessarily lead to bankruptcy. Closely related, Ward and Foster (1997) point out that studying only bankruptcy leads to an important bias because companies usually get into a financial distress cycle, lack financial flexibility and incur serious financial distress costs several years before going bankrupt. Additionally, Altman (1984a) has already highlighted the importance of using a definition of financial crisis regardless of its final outcome. We have thus used a financial criterion when defining a crisis, particularly because definitions of financial distress based on the company's failure to face its financial obligations are coherent with our ex-ante approach, which considers that financial distress costs are not limited to bankruptcy, as pointed out by Clark and Weinstein (1983).

Specifically, following Wruck (1990), Asquith et al. (1994), Andrade and Kaplan (1998) and Whitaker (1999), we adopt a definition of financial distress that evaluates the company's capacity to satisfy its financial obligations. We thus classify a company as financially distressed whenever its EBITDA is lower than its financial expenses, as this leads the firm to a situation in which it is not able to comply with its financial obligations. This criterion allows us to divide the samples into two groups and to construct a binary dependent variable that takes value one for financially distressed companies, and zero otherwise. According to Wood and Piece (1987), who questioned the accuracy of ex-post models when predicting financial distress ex-ante, we follow an ex-ante approach and, consequently, our definition of the dependent variable based on the condition of the EBITDA being lower than financial expenses is the most appropriate one.

3.2. A theoretically-based selection of explanatory variables

As we have already mentioned, the traditional selection of independent variables through sequential processes of eliminating them till reaching a maximum capacity of prediction leads to problems of overadjustment and contra-intuitive results. To avoid these problems we follow Pindado and Rodrigues (2004), who suggest a parsimonious selection of the independent variables. In fact, reviewing previous discriminant models that are still used in several countries (see Appendix 1), we can conclude that this kind of model does not require a huge set of variables in order to reach its maximum level of efficiency.

Accordingly, three theoretical underpinned variables have been chosen to enter our models, namely Earnings Before Interests and Taxes (EBIT), Financial Expenses (FE) and Retained Earnings (RE), all scaled by the replacement value of total assets at the beginning of the period instead of the book value[4].

First, we take into account that profitability ratios are typically used as measures of firm performance. We have chosen the EBIT variable, with the particularity of being scaled by a measure of the replacement value of total assets ($EBIT_{it}/K_{it-1}$), as a profitability indicator, since it captures the capacity of the firm to efficiently manage its assets and to obtain profitability, generating enough funds to face its financial obligations. In fact, creditors always use earnings to estimate the return generated by the

[4] A detailed description of all the variables used in this analysis can be found in Appendix 2. Table 2 reports summary statistics of these variables.

firm on borrowed capital (Claessens et al., 2003). Therefore, we expect this variable to negatively influence the PFD.

Second, research on PFD reveals the advantages of using a variable that considers the flow of financial expenses instead of the stock of debt. In fact, since the revision of the Z-score made by Altman et al. (1977) many other subsequent studies, such as Andrade and Kaplan (1998), call attention to debt variables being less an explanatory variable of financial distress than those variables of financial expenses. Asquith et al. (1994) also show how the leverage effect can be absorbed by the financial expenses effect, which constitutes one of the most frequent causes of financial distress in addition to the individual and sectoral components of economic crisis. In fact, Mackie-Mason (1990) already mentioned the advantage of using a measure of financial distress that does not include debt as an explanatory variable[5]. Additionally, Begley et al. (1996) point out that since the 1980s, firms have been continuously increasing their debt levels, without this having necessarily increased their probability of distress. This trend would explain the declining performance of the models proposed by Altman (1968) and Ohlson (1980). In short, recent literature shows that the flow of financial expenses imposes stricter limits on the company's policies than the stock of debt. Therefore, we include the variable of financial expenses scaled by the replacement value of total assets (FE_{it}/K_{it-1}) in our model, expecting a positive relation with the probability of financial distress.

Third, retained earnings (RE) are the total amount of reinvested earnings and/or losses of a firm over its entire life. This is a measure of cumulative profitability over time that remains one of the most crucial predictors of financial crisis. Particularly,

[5] The studies by Graham (1996) and Leary and Roberts (2004) use this suggestion in their capital structure models.

Routledge and Gadenne (2000) highlight the usefulness of past profitability in predicting future results and capacity for self-financing. Moreover, Mayer (1990) concludes that retained earnings constitute a privileged source of funds for companies in eight countries, including the US, the UK and Germany. We have thus introduced this ratio of cumulative profitability to replacement value of total assets (RE_{it}/K_{it-1}) into our model. This ratio has a less straightforward interpretation than the two prior variables concerning its relation to financial distress probability. On the one hand, the pecking order theory proposed by Myers and Majluf (1984) highlights the company's preference for internal funds, suggesting a negative relation between cumulative profitability and the PFD. On the other hand, Dhumale (1998) argues that the behaviour of this ratio can test for Jensen's (1986) free cash flow theory. According to this theory, the financial distress probability would increase as the company's cumulative profitability rises, since the availability of internal funds under management control could lead companies, especially those with less valuable investment opportunities, to misuse these funds by undertaking negative net present value projects. In this case, the expected relationship between the PFD and cumulative profitability would be positive.

3.3. Econometric specification

This study proposes a model to obtain a PFD that includes the above-described variables. We use explanatory stock variables evaluated at the beginning of the period and flow variables of the period, as suggested by Cleary (1999).

Given an objectively obtained binary dependent variable, the logistic regression technique determines the extent to which a set of variables containing useful information are able to classify every unit (every firm in our sample) in one category or the other (financially distressed or non-financially distressed). The logistic regression is expressed in terms of the odds ratio, which relates the probability of the event occurring to the probability of the event not occurring, as follows:

$$\log\left(\frac{\text{Prob(event)}}{\text{Prob(noevent)}}\right) = \beta_{0+}\beta_1 x_1 + \ldots + \beta_n x_n \qquad (1)$$

In our study, the odds ratio concerns the probability of being financially distressed according to the criterion described in Section 3.1. The independent variables in our logistic regression, whose effect on the probability of the firm being financially distressed was theoretically justified in Section 3.2, are: Earnings Before Interests and Taxes (*EBIT*), Financial Expenses (*FE*) and Retained Earnings (*RE*).

Accordingly, our logistic model to estimate the PDF is as follows:

$$\log\left(\frac{\text{Prob(event)}}{\text{Prob(noevent)}}\right) = \beta_0 + \beta_1 EBIT_{it}/K_{it-1} + \beta_2\, FE_{it}/K_{it-1} + \beta_3\, RE_{it-1}/K_{it-1} + d_t + \eta_i + u_{it} \qquad (2)$$

where all variables are indexed by an i for the individual cross-sectional unit (i= 1,…,N) and a t for the time period (t= 1,…, T). Additionally, d_t is the time effect, η_i denotes the individual effect, and the random disturbance is u_{it}~IID(0,σ^2).

The coefficients to be interpreted are β_1, β_2 and β_3[6]. The first (β_1) is associated with the capability of assets to generate returns, and is thus expected to be negative. The second (β_2) is predicted to be positive, since we expect the PFD to increase as the company's risk of not being able to comply with its financial obligations rises. Finally, the third coefficient (β_3) is expected to be negative if the economic agents' expectations are based on past profitability and self-financing, and positive when a large amount of

[6] It is worthwhile noting that, although the marginal effect that any independent variable in the regression has on the probability of the firm being financially distressed does not come directly from the beta coefficients, the interpretation of the signs of these coefficients is similar to that of the ordinary least squares regression.

internal funds leads to inefficiencies.

4. STRATEGY AND METHODS OF ESTIMATION

In this section, we describe the strategy implemented in order to obtain an indicator of the PFD. Figure 1 portrays the several stages followed in our new approach. The method of estimation of our models is also discussed in this section.

4.1. Strategy of estimation

Our strategy consists firstly of developing the econometric specification of the model according to financial theory, as has already been described in the previous section.

In a second stage, our study presents the innovation of estimating this model by using panel data methodology. Specifically, we estimate panel data models with a discrete dependent variable, since this methodology allows us to verify the significance of the model coefficients through the estimation of fixed and random effect panel data models that are robust to unobservable heterogeneity. Note that the implementation of this second stage requires the selection of a sample that allows us to work with data panels in which companies are chosen according to their financial distress situation in each period. Additionally, our concept of financial distress is compatible with large data panels, allowing us to use panel data methodology in order to consistently estimate the models of financial distress probability.

However, our PFD does not stem from these panel data models, since they eliminate unobservable heterogeneity. In other words, the panel data estimation removes the individual effects from the error term and, consequently, it does not account for the firm-specific contribution to the prediction of the probability of financial distress. The third stage in our approach addresses this issue. Specifically, once the robustness and the correct specification of the model have been tested for, we estimate a cross-sectional regression for each year, thus obtaining a good indicator of the PFD for each company and year.

4.2. Methods of estimation

Logit analysis is an appropriate explanatory technique for our study since our dependent variable is a binary variable. The research carried out in the 1980s consolidated logit analysis as a better estimation methodology than discriminant analysis, since the hypotheses on which the latter relies do not generally hold[7]. Consequently, we prefer to use logit analysis instead of discriminant analysis for several reasons. First, as discussed by Karel and Prakash (1987), discriminant analysis requires strict multivariate normality and homoskedasticity across groups, whereas logit analysis does not strictly require these assumptions. Second, logit analysis is often preferred even when these assumptions hold, mainly because of its ability to incorporate non-linear effects, and because of other technical features (Hair et al., 1995). Finally, discriminant analysis is not suitable for dealing with unobservable heterogeneity and other characteristics common in panel data samples.

Panel data models for discrete dependent variables allow us to correct the specification of the model by eliminating the bias of omitted variables that arises when

[7] As pointed out by Hair et al. (1995), the hypotheses of homoskedasticity of variances and covariances matrices and of multinormality of the variables do not generally hold.

the specific unobservable effects (η_i) are correlated with the explanatory variables in non-linear models. To this respect, it is necessary to distinguish between fixed effect models – those in which a relationship between the specific effect and the remaining right-hand side variables is not assumed – and random effect models – those in which this relation is functionally specified.

Regarding fixed effect models, the conditional likelihood estimator proposed by Chamberlain (1980), when feasible, allows us to obtain consistent estimates of the parameters in the presence of individual effects that are no longer dependent on the specific effect. On the other hand, Chamberlain (1984) proposes a random effect estimator that specifies the conditional distribution of η_i on explanatory variables. Specifically, this procedure is based on a parameterization of the correlation between the unobservable effect and the regressors, in such a way that the latter are considered time-variant explanatory variables of the former and of a random time-invariant term.

The choice between fixed and random effect models depends on the characteristics of the explanatory variables. On the one hand, when all the explanatory variables are expected to be strictly exogenous, the fixed effect model would yield good results if the estimation sample (observations for which there is a change in the regime between sample periods) is large enough regarding sample size, and there is temporal variation in the explanatory variables in order to identify the individual effect. On the other hand, Arellano and Honore (1999) highlight that the random effect model works better if explanatory variables are not strictly exogenous, samples show insufficient changes, or the contribution to the maximum likelihood function by the variation in explanatory variables is not enough.

Therefore, the preference for one of these models basically depends on the assumptions about the dependence of the error distribution on the explanatory variables.

Given the difficulty in establishing this relation, we follow Arellano and Honore (1999) in suggesting the convenience of estimating both models.

Nonetheless, although these models provide robust estimates of the parameters, they do not allow us to directly obtain a PFD because they do not take into account the individual effects. Overcoming this limitation can only be indirectly obtained by cross-sectionally estimating the PFD for each year.

5. RESULTS

In this section, we first present the estimation results of the random and fixed effect logistic regression models. We then tabulate the main statistics of the estimated probabilities as well as the percentages of correct classification produced by cross-section models that have been estimated for all the countries and years. Note that the same set of variables we proposed in Section 3 is always used when estimating such models.

5.1. Estimation results of the panel data models

Tables 3 and 4 present the results for the fixed and random effect models, respectively. The goodness of fit tests point to the high explanatory power of all the variables in both the fixed effect models (see likelihood ratios, LR, in Table 3) and in the random effect models (see Wald tests in Table 4). Additionally, Wald tests of the joint significance of the time dummies are presented, which validate the use of such variables in both models and for all countries, thus confirming, as was expected, that there have been fluctuations in the financial distress processes over time. These results show that the consideration of these dummy variables is important, since they allow us to accommodate the impact of changes in the macroeconomic environment.

Finally, the estimation of random effect models includes additional tests that verify the existence of unobservable heterogeneity. As shown in Table 4, the additional panel-level variance component is parameterized as $\ln(\sigma^2_\eta)$. The standard deviation, σ_η, is also reported in Table 4, and it is used to obtain a third indicator, (Rho)$\rho = \sigma^2_\eta / (\sigma^2_\eta+1)$, which is the proportion of the total variance contributed by the panel-level variance component. When Rho is approximately zero, the panel-level variance component is unimportant and the panel estimator is not different from the pooled estimator (logit estimator). In our study, the null hypothesis of equality of both estimators is rejected, and the existence of unobservable heterogeneity is thus confirmed. According to these results, we conclude that the proposed model needs to be validated by using a panel data methodology in order to control for unobservable heterogeneity.

We now turn our attention to the estimated coefficients in our panel data models. As they are all statistically significant and of the theoretically-expected sign for all the countries and using both methodologies, we are going to make a joint description of the results.

First, the variable that captures profitability ($EBIT_{it}/K_{it-1}$) negatively affects the PFD. This evidence is consistent with all the studies referred to in Appendix 1. Second, the effect of financial expenses (FE_{it}/K_{it-1}) is positive, which confirms our expectations about the capacity of this variable to capture the firm's financial vulnerability. The significance of the coefficient on the financial expenses variable is similar to that obtained in prior studies, such as Altman, Haldeman and Narayanan (1977), and Begley et al. (1996). Finally, the coefficient on cumulative profitability (RE_{it}/K_{it-1}) is negative for the US and the UK, which confirms the consequences of past profitability in determining the firm's financial structure (Opler and Titman, 1996) and supports our pecking order hypothesis. In contrast, the sign on this cumulative profitability variable for Germany is positive. Following the arguments in Dhumale (1998), this result may be interpreted as evidence of the free cash flow theory. Alternatively, it may be due to the role of banks leading German industrial groups, as has been justified by Hoshi et al. (1991) for the Japanese case.

According to these results, we can conclude that the PFD is explained in essence by the company's efficiency in extracting returns from its assets, and by the trade-off between this way of generating funds and the need to comply with its financial expenses during the financial year. We also find that, in general, higher historical profitability tends to reduce the company's PFD, which can serve as a cushion to provide wider financial solutions to the crisis. Alternatively, we also find that the decision-making process can be dangerously delayed when banks support the company's financing, as occurs in Germany.

Overall, we must highlight the relevance of our results mainly because, being similar for all the countries in both types of models, they constitute a strong validation of our approach. In fact, this evidence confirms that it is possible to build a more parsimonious model leading to a general and stable indicator of the PFD that can be used in different economic and legal contexts.

5.2. Estimation results of the probability of financial distress

We have already checked that our panel data model is correctly specified and that the variables used to explain the PFD are validated and supported by financial theory. The next step is to cross-sectionally estimate this correctly specified model for each year and country in order to obtain an indicator of such a probability that takes into account the firm individual effect.

The relevant output from our models is the probability of financial distress obtained for each company and year. Consequently, Table 5 provides the summary statistics of the estimated probabilities of financial distress. As can be seen in the table, the mean of the probability of financial distress is quite low (around 0.08), since our sample does not comprise only distressed firms. This result supports the goodness of our ex-ante estimation of the probability of financial distress. Additionally, there are some differences in the mean values across countries, which may be interpreted as the consequence of the different institutional context. Another indicator supporting the goodness of our approach is the small standard deviation showed by the probability of financial distress in all countries.

The percentage of correct classification also supports our approach. As can be seen in Table 5 the percentage of correct classification is quite stable across countries and years, its mean value is 95.20, and it is always above 90 percent. In any case, our discussion is not focussed on the percentage of correct classification, since our main concern is not to predict financial distress but to offer a model of its probability (Palepu, 1986)[8].

Finally, we highlight the importance of obtaining consistent estimates of the

---

[8] In fact, this percentage depends on the cutoff point, and the most common criticism relies on the fact that this point is usually determined ex-post, by a process of trial and error, without taking into account the fact that the probability of failure for the sample is not the same as that of the population. This process of classification can be particularly misleading when the loss functions of the errors are quite asymmetrical (see Hsieh, 1993) and, consequently, maximizing the percentage of correct classification can be quite different from minimizing the total error costs.

probabilities of financial distress per company and year, since they constitute necessary inputs for financial models that need to incorporate a measure of ex-ante financial distress costs. Therefore, the works by Opler and Titman (1994), Andrade and Kaplan (1998) or Dichev (1998) can be extended by making use of a concept of ex-ante financial distress costs as the product of the probability of financial distress and the ex-post financial distress cost perceived by investors. In this way, our approach is a step forward, since it allows researchers to obtain a good measure of the probability of financial distress.

6. CONCLUSIONS

This paper offers a new approach to estimating the probability of financial distress which can be applied to different economic and legal contexts. To achieve this aim, we have first developed a theoretically supported model relying on a financial criterion of financial distress that is independent of legal institutions. Additionally, independent variables have been selected according to financial theory in order to avoid the traditional criticism concerning the lack of financial justification of the results. The chosen variables are earnings before interest and taxes, financial expenses and retained earnings.

We have then tested for the specification of the resultant logistic regression by using panel data methodology in order to eliminate the unobservable heterogeneity. The results obtained confirm the specification of the model, and reveal that all the coefficients are statistically significant and of the expected sign for all the countries, using both fixed and random methodologies. Specifically, we find that the probability of financial distress is accurately explained by the company's returns on assets, and the consequent trade-off between this way of generating funds and the company's need to comply with its financial expenses during the financial year. This evidence strongly validates our approach to estimating the probability of financial distress, giving rise to an indicator that can be used in different economic and legal contexts.

The need to incorporate the specificity of each company into the estimates of the PFD has motivated a cross-sectional estimation of this correctly specified model. The results definitely confirm the accuracy of our model, and show a very high percentage of correct classification for all countries and years. Furthermore, we provide important results concerning the estimation of a consistent and theoretically based probability of financial distress per company and year that is fundamental for developing models of financial distress costs.

Finally, it is worthwhile remarking that our approach is a step forward in the literature, since it provides a good measure of the probability of financial distress that facilitates the calculation of ex-ante financial distress costs as the product of such a probability and the ex-post financial distress costs.

Appendix 1: Comparison of some relevant models developed in Germany, the US and the UK

Country	**US**	**US**	**US**	**US**	**UK**	**Germany***
Author	Altman, E.	Altman, E. , Haldeman, R. Narayanan, P.	Ohlson, J.	Zmijewski, M.	Taffler, R.	Deustch Bundesbank
Year of the study	1968	1977	1980	1984	1984	1992
Estimation (LDA= Linear Discriminant Analysis)	LDA	LDA	LOGIT	PROBIT	LDA	LDA
Samples (D =Distressed; N = Normal)	33 D, 33 N	53 D, 58 N	105 D, 2058 N	129 D, 2241 N	46 D, 46 N	677 D, 677 N
Sample period	1946-1965	1962-1975	1970-1976	1972-1978	1968-1976	1987-1992
Overall prediction accuracy (estimation sample)	T-1 : 95 T-2 : 72 T-3 : 48 T-4 : 29 T-5 : 36	T-1 : 92,8 T-2 : 89 T-3 : 83,5 T-4 : 79,8 T-5 : 76,8	T-1 : 96,12 T-2 : 95,55 T-3 : 92, 84 and other results for other cut-off points	Estimations for different percentages of insolvent firms for T-1	T-1: Distressed: 95.4 Normal:1 00	T-1: 89,25
Number of ratios	5	7	9	3	4	3
Liquidity and Working Capital	1	1	2	1	3	-
Profitability and Cash Flow	2	3	4	1	1	2
Financial Structure and Debt Service	1	2	2	1	-	1
Activity	1	1	1	-	-	-

*German data were obtained from the proceedings of the Paris Meeting in May 1998 of the European Committee of Central Balance Sheet Data Offices, particularly from the presentation of the Working group nº1: Credit Risk Analysis, Applications in some European Central Banks and Related Organizations.

Appendix 2: Variable definitions

- *Replacement value of total assets*

$$K_{it}=RF_{it}+(TA_{it} - BF_{it}) \tag{A.1}$$

where RF_{it} is the replacement value of tangible fixed assets, TA_{it} is the book value of total assets, and BF_{it} is the book value of tangible fixed assets. The last two terms were obtained from the firm's balance sheet, and the first was calculated according to Perfect and Wiles (1994):

$$RF_{it} = RF_{i,t-1}\left[\frac{1+\phi_t}{1+\delta_{it}}\right]+I_{it} \tag{A.2}$$

for $t>t_0$ and $RF_{it_0}=BF_{it_0}$, where t_0 is the first year of the chosen period, in our case 1990. On the other hand, $\delta_{it}=D_{it}/BF_{it}$ and $\phi_t=(GCGP_t-GCGP_{t-1})/GCGP_{t-1}$, where $GCGP_t$ is the growth of capital goods prices reported in the Main Economic Indicators published by the Organization for Economic Cooperation and Development (OECD).

- *Earnings before interests, taxes and amortization*: $EBITDA_{it}=RVT_{it}-XOPR_{it}$

where RVT_{it} is the total revenue, and $XOPR_{it}$ denotes operating expenses.

- *Return ratio:* $EBIT_{it}/K_{it-1}$, where $EBIT_{it}$ is earnings before interests and taxes.

- *Financial Expenses ratio:* FE_{it}/K_{it-1}, where FE_{it} denotes financial expenses.

- *Cumulative Profitability ratio:* RE_{it-1}/K_{it-1}, where RE_{it} denotes retained earnings.

REFERENCES

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

Altman E. (1984a), 'A further empirical investigation of the bankruptcy cost question', *The Journal of Finance*, Vol. 39, pp. 1067-1089.

Altman, E. (1984b), 'The success of business failure prediction models: An international survey', *Journal of Banking & Finance*, Vol. 8, pp. 171-198.

Altman, E., R. Haldeman and P. Narayanan (1977), 'Zeta analysis: A new model to identify bankruptcy risk of corporation', *Journal of Banking & Finance*, Vol. 1, pp. 29-54.

Altman, E. and A. Saunders (1997), 'Credit risk measurement: developments over the last 20 years', *Journal of Banking & Finance*, Vol. 21, pp. 1721-1748.

Altman, E. and P. Narayanan (1997), 'An international survey of business failure classification models', *Financial Markets, Institutions and Instruments*, Vol. 6, pp. 1-56.

Andrade, G. and S. Kaplan (1998), 'How costly is financial, (not economic) financial distress – Evidence from highly leveraged transactions that became distressed', *The Journal of Finance*, Vo. 53, pp. 1443-1493.

Arellano, M. and B. Honore (1999), 'Panel data models: Some recent evidence', Working Paper, Princeton University, http://www.princeton.edu/~honore/papers).

Asquith, P., R. Gertner and D. Scharfstein (1994), 'Anatomy of financial distress: An examination of junk-bond issuers', *Quarterly Journal of Economics*, Vol. 109, pp. 625 - 658.

Barnes, P. (1987), 'The analysis and use of financial ratios: a review article', *Journal of*

Business Finance & Accounting, Vol. 14, pp. 449-461.

Barnes, P. (1990), 'The prediction of takeover targets in the U.K. by means of multiple discriminant analysis', *Journal of Business Finance & Accounting* Vol. 17, pp. 73-83.

Beaver, W. (1966), 'Financial ratios as predictors of failure', *Journal of Accounting Research Supplement, Empirical research in accounting: selected studies*, Vol. 4, pp. 71-111.

Begley, J., J. Ming and S. Watts (1996), 'Bankruptcy classification errors in the 1980s: An empirical analysis of Altman and Ohlson's models', *Review of Accounting Studies,* Vol. 1, pp. 267-284.

Chamberlain, G. (1980), 'Analysis of covariance with qualitative data', *Review of Economic Studies*, Vol. 47, pp. 225-238.

Chamberlain, G. (1984), 'Panel data', in Griliches, Z. and M.D. Intriligator, (Eds.), Handbook of Econometrics, Vol. 2, North -Holland, Amsterdam.

Clark, T. and M. Weinstein (1983), 'The behaviour of the common stock of bankrupt firms', *The Journal of Finance*, Vol. 38, pp. 489-504.

Claessens, S., S. Djankov and L. Klapper (2003), 'Resolution of corporate distress in East Asia', *Journal of Empirical Finance*, Vol. 10, pp. 199-216.

Claessens, S. and L. Klapper (2002), 'Bankruptcy around the world: Explanations of its relative use', Policy Research Working Paper World Bank, 2865.

Cleary, S. (1999), 'The relationship between firm investment and financial status', *The Journal of Finance*, Vol. 54, pp. 673-692.

Dhumale, D. (1998), 'Earnings retention as a specification mechanism in logistic bankruptcy models: A test of the free cash flow theory', *Journal of Business Finance & Accounting*, Vol. 25, pp. 1005-1023.

Dichev, I. (1998), 'Is the risk of bankruptcy a systematic risk?', *The Journal of Finance*, Vol 53, pp. 1131-1147.

Graham, J. R. (1996), 'Debt and the marginal tax rate', *Journal of Financial Economics*, Vol. 41, pp. 41-73

Hair, J., R. Anderson, R. Tatham and W. Black (1995), Multivariate data analysis with readings (Prentice Hall).

Hoshi, T., A. Kashyap and A. Scharfstein (1991), 'Corporate Structure Liquidity and Investment: Evidence from Japanese Industrial Groups', *The Quarterly Journal of Economics*, Vol. 106, pp. 33-60.

Hsieh, S. (1993), 'A note on the optimal cutoff point in bankruptcy prediction models', *Journal of Business Finance & Accounting*, Vol. 20, pp. 457-464.

Jensen, M. (1986), 'Agency costs of free cash flow, corporate finance and takeovers', *The American Economic Review*, Vol. 76, pp. 323-329.

Jones F. (1987), 'Current techniques in bankruptcy prediction', *Journal of Accounting Literature*, Vo. 6, pp. 131-164.

Karels, V. and A. Prakash (1987), 'Multivariate normality and the forecasting of business bankruptcy', *Journal of Business Finance & Accounting*, Vol. 14, pp. 573-593.

Leary, M. and M. Roberts (2004), 'Do firms rebalance their capital structures?' Duke Fuqua School of Business, 1/1/2004.

Mackie-Mason, P. (1990), 'Do taxes affect corporate financing decisions?', *The Journal of Finance*, Vol. 44, pp. 1471-1493.

Mayer, C. (1990), 'Financing systems, corporate finance, and economic development', in R.G. Hubbard, (Eds.), Asymmetric Information, Corporate Finance and Investment. University of Chicago Press, Chapter 12.

Myers, S. and S. Majluf (1984), ‘Corporate financing and investment decisions when firms have information that investors do not have’, *Journal of Financial Economics*, Vol. 13, pp. 187-221.

Ohlson, J. (1980), ‘Financial ratios and the probabilistic prediction of bankruptcy’, *Journal of Accounting Research*, Vol. 18, pp. 109-131.

Opler, T. and S. Titman (1994), ‘Financial distress and corporate performance’, *The Journal of Finance*, Vol. 49, pp. 1015-1040.

Opler, T. and S. Titman (1995), ‘Financial distress and capital structure choice’, Unpublished manuscript, Ohio State University.

Palepu K. (1986), ‘Predicting takeover targets: A methodological and empirical Analysis’, *Journal of Accounting and Economics*, Vol. 8**,** pp. 3-35.

Perfect, S. and K. Wiles (1994), ‘Alternative constructions of Tobin's q: An empirical comparison’, *Journal of Empirical Finance*, Vol. 1, pp. 313-341.

Pindado J. and L. Rodrigues (2004), ‘Parsimonious models of financial distress in small companies’, *Small Business Economics*, Vol. 22, pp. 51-56.

Routledge J. and D. Gadenne (2000), ‘Financial distress, reorganization and corporate performance’, *Accounting and Finance*, Vol. 40, pp. 233-260.

Scott, J. (1981), ‘The probability of bankruptcy: A comparison of empirical predictions and theoretic models’, *Journal of Banking & Finance*, Vol. 5, pp. 317-344

Taffler, R. (1983), ‘The assessment of company solvency and performance using a statistical model’, *Accounting and Business Research*, Vol. 13, pp. 295-306.

Taffler, R. (1984), ‘Empirical models for the monitoring of the UK corporations’, *Journal of Banking & Finance*, Vol. 8, pp. 199-227.

Ward, T. and B. Foster (1997), ‘A note on selecting a response measure for financial distress’, *Journal of Business Finance & Accounting*, Vol. 24, pp. 869-879.

Whitaker, R. (1999), 'The early stages of financial distress', *Journal of Economics and Finance*, Vol. 23, pp. 123-133.

White, M. J. (1996), 'The costs of corporate bankruptcy: A US-European comparison' in Bhandari, J. S. and L.A. Weiss, Ed., Corporate bankruptcy, economic and legal perspectives, Cambridge University Press.

Wood, D. and J. Piesse (1987), 'The information value of MDA based financial indicators', *Journal of Business Finance & Accounting*, Vol. 14, pp. 27-38.

Wruck, K. (1990), 'Financial distress, reorganisation, and organisational efficiency', *Journal of Financial Economics*, Vol. 27, pp. 419-444.

Zavgren, C. (1985), 'Assessing the vulnerability to failure of American industrial firms: A logistical analysis', *Journal of Business Finance & Accounting*, Vol. 12, pp. 19-45.

Zmijewski, M. (1984), 'Methodological issues related to the estimation of financial distress prediction models', *Journal of Accounting Research Supplement*, Vol. 22, pp. 59-86.

Table 1
Number of companies and annual observations per country

Country	**Number of firms**	**Number of annual observations**					**Total number of observations**
		6	**7**	**8**	**9**	**10**	
Germany	186	102	168	80	117	1220	1687
US	1704	1014	1246	1416	1584	10040	15300
UK	491	246	406	544	639	2530	4365
Total	2381	1362	1820	2040	2340	13790	21352

Table 2
Summary statistics

	$EBIT_{it-1} / K_{it-1}$				FE_{it} / K_{it-1}				RE_{it-1} / K_{it-1}			
	Germany	US	UK	Total	Germany	UK	US	Total	Germany	UK	US	Total
Observations	1687	15300	4365	21352	1687	4365	15300	21352	1687	4365	15300	21352
Mean	0.0644	0.0641	0.0881	0.0690	0.0170	0.01718	0.02227	0.02081	0.0531	0.1682	0.0709	0.8940
Std. deviation	0.0710	0.1496	0.1145	0.1386	0.0122	0.01512	0.02367	0.02158	0.0738	0.2673	0.8912	0.7653
Minimum	-0.5102	-3.4922	-1.4984	-3.4922	0.000	0.000	0.000	0.000	-0.3443	-5.0947	-5.1747	-5.1747
Maximum	0.3526	1.155	0.5635	1.1550	0.0702	0.17603	0.78729	0.78729	0.4225	1.0404	1.2099	1.2099

Table 3
Results of the fixed effect logistic model by country

Country	Germany	US	UK
Observations	1501	13596	3874
Companies	186	1704	491
$EBIT_{it}$/ K_{it-1}	-50.9162*	-21.0379	-22.9463*
	-8.365768	0.94988	-2.156759
FE_{it}/ K_{it-1}	99.74603*	28.77826	58.83337*
	-35.47586	4.806255	-16.56535
RE_{it-1} / K_{it-1}	-3.80527	-0.28918***	-2.05432**
	-4.605058	0.151331	-0.960036
RE_{it-1} / K_{it-1}	-3.80527	-0.28918***	-2.05432**
	-4.605058	0.151331	-0.960036
Time χ^2 (8)[a]	16.18**	16.08**	2.01
LR χ^2 (11)[b]	198.74*	1482.98*	398.54*

*, **, *** indicate significance at 1, 5 and 10%, respectively.

[a] Wald test of the joint significance of the time dummy variables, asymptotically distributed as $\chi 2$ under the null of no relationship; degrees of freedom in parentheses.

[b] LR Maximum likelihood ratio test of goodness of fit, asymptotically distributed as $\chi 2$ under the null of no joint significance of the coefficients; degrees of freedom in parentheses.

Table 4
Results of the random effect logistic model by country

Country	**Germany**	**US**	**UK**
Observations	1501	13596	3874
Companies	186	1704	491
$EBIT_{it}/ K_{it-1}$	-49.8716* -5.594072	-25.069* -0.804619	-24.8538* -1.682444
FE_{it}/ K_{it-1}	90.01008* -20.68682	27.37314* -3.213095	57.15494* -10.45557
RE_{it-1} / K_{it-1}	-3.79571 -3.26847	-0.80705* -0.096274	-2.36303* -0.489623
Constant	-3.81391* -0.804423	-3.30518* -0.210062	-3.82889* -0.47621
$Ln\sigma^2_{\eta}$	1.413665 -0.487211	1.331225 -0.113193	0.960999 -0.317973
σ_{η}	2.027559 -0.493925	1.945682 -0.110119	1.616881 -0.257062
Rho	0.804343 -0.076675	0.791043 -0.01871	0.723322) -0.063635
Rho = 0 χ^2 (1)[a]	30.89*	461.66*	39.5*
Time χ^2 (8)[b]	15.19***	21.05*	2.36
Wald χ^2 (11)[c]	82.88*	1019.7*	228.27*

*, **, *** indicate significance at 1, 5 and 10%, respectively.

[a] Test of the joint significance of individual effects.

[b] Wald test of the joint significance of the time dummy variables; asymptotically distributed as $\chi2$ under the null of no relationship; degrees of freedom in parentheses.

[c] Wald test of goodness of fit, asymptotically distributed as $\chi2$ under the null of no joint significance of the coefficients; degrees of freedom in parentheses.

Table 5
Summary statistics of estimated probabilities

Statistic	Mean			Standard deviation			Percentage of correct classifications		
Country	Germany	US	UK	Germany	US	UK	Germany	US	UK
Probability 91	0.007	0.136	0.047	0.01	0.234	0.176	99.29	91.31	98.17
Probability 92	0.042	0.109	0.067	0.153	0.239	0.174	96.23	94.53	95.72
Probability 93	0.114	0.102	0.070	0.196	0.227	0.183	91.95	94.05	95.4
Probability 94	0.109	0.100	0.066	0.276	0.222	0.178	96.55	94.08	95.98
Probability 95	0.091	0.104	0.067	0.205	0.230	0.195	93.01	94.54	97.15
Probability 96	0.087	0.098	0.063	0.210	0.228	0.181	94.02	95.41	96.24
Probability 97	0.082	0.101	0.065	0.223	0.227	0.201	97.25	94.49	96.21
Probability 98	0.051	0.115	0.080	0.119	0.236	0.207	94.35	92.85	96.25
Probability 99	0.067	0.116	0.067	0.196	0.232	0.162	95.76	94.42	93.87
Global	0.074	0.108	0.066	0.195	0.230	0.182	95.336	93.998	96.205

Figure 1

Stages of the new approach to estimating the probability of financial distress