# Forecasting Distress in European SME Portfolios

Ferreira Filipe, Sara and Grammatikos, Theoharry and Michala, Dimitra

University of Luxembourg, University of Luxembourg, University of Luxembourg

1 February 2014

# Forecasting Distress in European SME Portfolios

Sara Ferreira Filipe*, Theoharry Grammatikos*, Dimitra Michala*

Luxembourg School of Finance

February 2014

## ABSTRACT

In the European Union, small and medium sized enterprises (SMEs) represent 99% of all businesses and contribute to more than half of the total value-added. In this paper, we develop distress prediction models for SMEs using a dataset from eight European countries over the period 2000-2009. We examine idiosyncratic and systematic covariates and find that the first discriminate between healthy and distressed firms based on their relative level of risk, whereas the second move the overall distress rates. Moreover, SMEs across Europe are vulnerable to the same idiosyncratic factors but systematic factors vary in different regions. Also, micro SMEs are more vulnerable to these systematic factors compared to larger SMEs.

The paper contributes to the literature in several ways. First, using a sample with many micro companies, it offers unique insights into the European small business sector. Second, it is the first paper to explore distress in a multi-country setting, allowing for regional comparisons and uncovering regional vulnerabilities. Third, by incorporating systematic dependencies, the models can capture changes in overall distress rates and comovements during economic cycles.

*JEL Codes*: C13, C41, C53, G33

*Keywords*: credit risk, distress, forecasting, SMEs, discrete time hazard model, multi-period logit model, duration analysis

---

* Luxembourg School of Finance, University of Luxembourg, 4 rue Albert Borschette, L-1246, Luxembourg. Emails: sara.ferreira@uni.lu; theoharry.grammatikos@uni.lu; dimitra.michala@uni.lu. Dimitra Michala is the corresponding author (Tel: +352 4666446805).

We would like to thank Jos Van Bommel, Michel Dietsch, Federico Galizia, Victoria Golikova, Jean-Daniel Guigou, Helmut Kraemer-Eis, Edward Olding, and participants in the 6th Annual EIBURS Meeting, the LSF Internal Seminar series, the International Conference on Economic and Social Development (Moscow Higher School of Economics), the 2nd Credit Analysis and Risk Management Conference (University of Basel), and the XXI Finance Forum (Instituto de Empresa Business School) for precious comments. Special thanks to Mounir Shal for help with the database. This research is financially supported by the European Investment Bank University Research Sponsorship Program. All errors are our own.

SMEs play a crucial role in most economies. In the Organization for Economic Cooperation and Development (OECD) countries, SMEs account for 95% of all enterprises and generate two-thirds of employment. In the European Union in particular, SMEs represent 99% of all enterprises and contribute to more than half of the value-added created by businesses. Despite their importance, SME credit risk remains largely unexplored by the academic literature, mainly due to their information opaqueness and lack of available data.

In this paper, we explore a unique dataset that allows new insights into the European SME sector and its credit risk characteristics. First, our sample includes a very high number of micro companies. This focus on the micro sector is important, as nine out of ten European SMEs are micro-enterprises.[1] Second, we include SMEs from eight European economies and examine distress[2] within regions, unlike earlier studies, which have focused on a single economy.[3] Third, we consider systematic factors, such as the macroeconomy, bank lending conditions, and legal aspects. Hence, we are able not only to compute individual distress probabilities, but also to estimate overall distress rates in the economy and capture distress comovements.

Our paper therefore contributes to the overall literature on corporate credit risk, and on SME risk in particular. It is well known that, unlike larger corporations with easier access to capital markets, SMEs offer more challenges in their credit risk modelling.[4] In fact, widely used structural market-based models, such as the distance-to-default measure inspired by Merton (1974), cannot be applied in the SME setting. Instead, empirical predictive models such as credit scoring approaches (i.e. Altman, 1968) are the most common. In the early credit rating literature, academics mostly use accounting ratios to predict firm distress.[5] The first credit scoring study that focuses on small businesses is Edminster (1972), who analyzes nineteen financial ratios and develops a model using multivariate

[1] Micro-enterprises have fewer than 10 employees and a turnover under €2million.
[2] We provide the definition of distress in subsection II.A. and the definition of SMEs in subsection II.B.
[3] For instance, Altman and Sabato (2007) focus on the US, Nam et al. (2008) on Korea, Altman et al. (2010) on the UK, and Jacobson et al. (2013) on Sweden.
[4] Dietsch and Petey (2004) explore such differences between SMEs and corporations.
[5] See Altman and Narayanan (1997) and Altman and Saunders (1998) for literature reviews.

discriminant analysis.

Recently the need for SME-specific research has become pressing, in particular with the implications of Basel's II special treatment requirements for SMEs exposures and the high numbers of distressed SMEs during the crisis. Berger and Udell (2006) point the need for lending technologies specifically designed for SMEs. Similarly, Beck et al. (2008) find that financing patterns have important differences for SMEs compared to large firms. In line with these new concerns, Altman and Sabato (2007) develop a one-year default prediction model for SMEs using only accounting information. They apply panel logit estimation on a sample of around 2,000 US firms over the period 1994-2002 and find that their model outperforms generic corporate models such as Altman's Z"-score (Altman and Hotchkiss, 2005).

Moreover, Grunert et al. (2005) and other authors have noted the possibility of using qualitative variables in default prediction models to improve discrimination. Stein (2002) is the first to investigate the importance of "soft" information in borrower-bank relationships. Specifically in the case of SMEs, where there is usually a problem of scarcity of reliable "hard" financial information, such non-financial elements can be very useful in distress prediction. DeYoung et al. (2008) find though that dependence on only "soft" information can increase default events, especially in situations that there is distance between the small business borrower and the bank. Altman et al. (2010) combine both qualitative and financial information in a default prediction model for SMEs. By applying multi-period logit estimation to a large sample of UK SMEs for the period 2000-2007, they find that data relating to legal action by creditors, company filings, and audit reports/opinions significantly increase the performance of their model. Such information though is not always available well in advance for timely predictions.

Another strand of literature, though not focusing on SMEs, analyses the additional benefit of using macroeconomic variables to forecast distress.[6] Wilson (1997a, 1997b) develops an aggregate credit risk model that explicitly links macroeconomic factors and corporate sector default rates.[7] Carling et al. (2007)

[6] For a nice literature review on the incorporation of systemic influences into risk measurements, see Allen and Saunders (2004).

[7] Jacobson et al. (2005) apply a similar framework to default risk of Swedish companies. Bruneau et al. (2012) follow the same approach to study firms' financial fragility in France.

find that the output gap, the yield curve and consumers' expectation add significant predictive power to distress models. Duffie et al. (2007) incorporate macroeconomic covariates to estimate conditional probabilities of corporate default for a sample of US listed industrial firms. Similarly, Campbell et al. (2008) introduce the macroeconomic environment through financial market variables. Nam et al. (2008) use specifically the volatility of the exchange rate and Koopman et al. (2009) condition on business cycle effects, bank lending conditions and financial market variables. Jacobson et al. (2013) use a large Swedish dataset and find that macroeconomic variables are important for explaining the time-varying default likelihoods. Authors have also noted the importance of industry effects - for instance, Chava and Jarrow (2004) observe improving forecasting performance by including industry groupings in their models.

In an early study, Berger and Udell (1998) discuss the impact of the macroeconomic environment on small firms. Some years later, Glennon and Nigro (2005) and Altman et al. (2010) are the first to examine business cycle effects on SMEs defaults in particular. Glennon and Nigro (2005), using a dataset of US loans guaranteed by the Small Business Administration Scheme, include business cycle dummy variables, the industrial production index growth and rates of regional business bankruptcies to capture regional and industry economic conditions on the default hazard. They find that the success or failure of a small loan is closely related to both regional and industrial economic conditions. Altman et al. (2010) use sector-level failure rates and also report a significant relationship with failure probability. Our study differs from previous ones since we examine a larger variety of systematic factors, ranging from exchange rates to bank lending conditions, and use a wider sample that includes SMEs from different European countries, allowing for regional models and comparisons.

We examine a large number of idiosyncratic and systematic variables and, in addition to determining the importance of profitability, coverage, leverage and cash flow indicators, we note that SMEs in urban areas and SMEs with less than three shareholders have higher distress probabilities. We also find that the exchange rate, the economic sentiment, the credit supply and the bankruptcy codes significantly affect overall distress rates in the European economy. Nevertheless, industry effects often do not demonstrate significance. Moreover, we examine interaction effects between SMEs' size and systematic variables and find that as SMEs become larger, they are less vulnerable to systematic factors. In terms of prediction power, idiosyncratic factors do a good job in discriminating between healthy and distressed firms based on their relative level of risk, whereas systematic factors are the most important determinants of the overall distress rates.

We then split our sample in regional groups and find that SMEs in different regions are vulnerable to the same idiosyncratic factors (related to profitability, coverage, leverage and cash flow) but coefficients' levels differ among these regions. SMEs in different regions are also exposed to different systematic factors, according to region-specific conditions and characteristics. These findings indicate the importance of using regional models for distress prediction in international portfolios. We also split our sample in four rolling window periods (each one lasting five years) and find that whereas sensitivities to idiosyncratic factors remain relatively stable over time, coefficients of systematic variables are more volatile, responding to changes in the prevailing macroeconomic conditions. Finally, we test the performance of our models with a battery of out-of-sample tests.

The paper is organized as follows: Section I describes the methodology and the reasons for its selection. Section II describes the dataset, discusses the choice of variables and presents summary statistics. Section III presents the models and discusses the estimation results as well as the robustness checks, and Section IV concludes.

## I. The Methodology

We follow Shumway (2001) and estimate the probability of distress over the next year using a multi-period logit model.[8] We assume that the marginal probability of distress (or hazard rate) over next year follows a logistic distribution and is given by:

[8] Shumway (2001) proves that a multi-period logit model is equivalent to a discrete-time hazard model with an adjusted standard error structure. He also shows that multi-period logit models are more appropriate to static ones for distress forecasting because they account for the fact that firms' financial conditions change through time (by using samples that include consecutive firm-year observations). In a static single-period model, one firm-year observation for each healthy firm is randomly selected from the available firm-years, whereas for distressed firms, the firm-year immediately prior to distress is (non-

$$h(t|x_{i,t-1}) = P(Y_{i,t} = 1|x_{i,t-1}) = \frac{1}{1+\exp(-\beta x_{i,t-1} - \gamma y_{t-1})}, \quad (1)$$

where $Y_{i,t}$ is an indicator that equals one if the firm is distressed in year $t$, $\beta x_{i,t-1}$ is a function of firm-specific characteristics that includes a vector of firm-specific variables $x_{i,t-1}$ known at the end of the previous year and $\gamma y_{t-1}$ is the baseline hazard function that includes some other time-dependent variables $y_{t-1}$. The baseline hazard influences similarly all firms in the economy and expresses the hazard rate in the absence of the firm-specific covariates $x_{i,t-1}$. In this paper, we follow Duffie et al. (2007), Campbell et al. (2008) and other authors and specify the baseline hazard using macroeconomic variables. A higher value of $\beta x_{i,t-1} + \gamma y_{t-1}$ implies a higher probability of distress.

However, test statistics produced by the logit program are incorrect because they assume that the number of independent observations is the number of firm-years and ignore the panel structure of the data. Calculating correct test statistics requires adjusting the sample size to account for dependence among firm-year observations. For this reason, we adjust the standard errors of our models for the number of firms in the samples (clustered - corrected standard errors).[9]

Finally, we account for the survivorship bias, which is the risk that SMEs are more likely to be in our sample if they are survivors, consequently, these firms have lower distress probabilities. Particularly in 2000, which is the first year of our sample period, all firms that are present in the database are survivors. This happens because 2000 is the year that our database becomes more complete. As firms enter the database later on, they are always survivors in the first year of their existence in the sample (firms that fail quickly simply are never included in the sample). Thus, we introduce one more factor, the "duration" variable that accounts for the "time-at-risk" of firms only during the sample period. This variable is the number of years that a firm stays in the sample and is measured in discrete time units. I.e., if an SME appears in the sample for three years in total, the value of this variable in the first year is one,

---

randomly) selected. This process introduces a bias. On the other hand, a multi-period logit model is estimated with data on each firm in each available year, as if each firm-year is a separate observation.

[9] Calculated from Huber/White sandwich covariance matrix; see Froot (1989), White (1994) and Wooldridge (2002).

in the second year two and in the third year three. By censoring the number of years that a firm existed before it joined the sample, we weight all firms on equal terms and account for duration dependence, since we allow the time a firm remains in the sample to directly affect the probability of distress, over and above its accounting data and the systematic factors.[10]

## II. The Data

In order to estimate the multi-period logit model, we need an indicator of distress (dependent variable) and a set of predictors (independent variables). We use the Amadeus and Orbis Europe databases (both available from Bureau Van Dijk) to detect the status of each firm in each year and extract the raw data that include financial and qualitative information. Finally, we use the European Statistical Service's (Eurostat), the European Central Bank's (ECB), the World Bank and Datastream databases for the systematic variables.

In this part, we first discuss the definition of distress that we adopt, we then explain what criteria need to be met for a company to be included in the sample and finally, we describe the examined predictive variables and the procedure we follow to select the best among them.

### A. Definition of Distress

Tracking the status of SMEs is a very challenging task. There are many reasons for which an SME can go out of business but owners rarely report these reasons and authorities rarely document them. Watson and Everett (1996) find that small businesses often close for reasons other than distress. For example, the owner may close the firm voluntarily to accept employment elsewhere or retire. Headd (2003) finds that only one third of start-ups close under conditions that owners consider unsuccessful. Even when SMEs are financially distressed, they often do not follow formal insolvency proceedings but simply drop out of business. Gilson and Vetsuypens (1993) find that, in the US, many insolvency filings are missing for

[10] The "duration" variable is still an imperfect measure though since we can underestimate the lifespan of firms that default in the beginning of the sample period.

distressed firms. As a result, when studying SMEs distress, it is important to distinguish between distress and closure. Below, we explain the distress definition that we use to overcome these difficulties.

We classify firm-years into two mutually exclusive categories: "healthy" and "distressed". A firm-year is distressed if the following two conditions are met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; and (ii) the firm appears with one of the following statuses - defaulted, in receivership, bankrupt, or in liquidation. If no status information is available, we consider only the cases where the firm leaving the sample has negative equity in its last year,[11] following Boss's (2010) definition of stock-based insolvency and Davydenko's (2012) definition of economic default.[12] A firm-year is "healthy" in all other cases (i.e. if a firm drops from the sample due to merger). We construct the distress indicator as follows: it equals one for the distressed firm-years and zero for the healthy ones.

Our aim here is to proxy for distress. Thus, we apply a definition of distress, beyond that determined only by legal insolvency procedures. As noted above, SMEs often do not follow such procedures. A characteristic example is Italy, where there is no established framework for SMEs to file for insolvency. Even in cases where there is such a framework, filings are not mandatory or they take a long period of time. Finally, when these procedures are mandatory, legal insolvency is often related to equity levels: In Germany, firms are obliged to file for bankruptcy once their equity turns negative (Davydenko and Frank, 2008).

[11] The Orbis and Amadeus databases cooperate in different countries with credit bureaus which provide firm status information. In some cases though, a firm leaves the sample but the status information remains outdated. It is only for these cases that we apply the negative equity assumption. We observe that negative equity is 200% more likely for firms that drop from the sample than for firms that remain active. Due to this assumption, we need to exclude year 2010 from the sample. As accounts for SMEs usually become available with a considerable time lag, accounts for 2011 were mostly missing in early 2012, when the database became available to us. As a result, we could not detect distressed firms for 2010, since we do not know which firms drop from the sample in the following year.

[12] Davydenko (2012) describes as economic default the point when a firm's equity turns negative and finds support for models in which the default timing is chosen endogenously. Boss et al. (2010) point out that the definition of distress has two bases: a stock-based insolvency and a flow-based insolvency. A stock-based insolvency occurs when a company has negative equity and a flow-based insolvency when a company's operating cashflow is insufficient to meet current obligations. Earnings before interest, taxes, depreciation and amortization (EBITDA) are often used in academic studies as a proxy for operating cashflow. In our sample, 61% of the companies that drop from the sample having negative equity, also report negative EBITDA for that year. For robustness purposes, we calculate overall distress rates assuming that firms that drop from the sample having negative EBITDA are distressed and find substantially similar results.

## B. Sample Selection

Our estimation sample consists of 2,721,861 firm-years observations (644,234 firms) out of which 49,355 are distressed. SMEs come from eight European countries, namely Czech Republic, France, Germany, Italy, Poland, Portugal, Spain and the United Kingdom. We keep a random one tenth of the firms from each country as a hold-out sample. The hold-out sample consists of 304,037 firm-year observations (71,823 firms) out of which 5,487 are distressed. We select the countries mostly due to data availability issues but also because they create a combination that nicely reflects the variability in the importance of SMEs across the EU. Table 1 provides an overview of the key indicators for SMEs in the EU27 and in the countries of our sample. As seen, in Italy, Portugal and Spain, SMEs have larger shares in employment and value added and higher density than in the EU on average. This suggests that SMEs in these economies have a more important role than in most EU countries. On the other hand, for France, Germany and the UK, these figures are always lower than the EU average. For Czech Republic and Poland, the employment share and value added of SMEs is similar. Czech Republic though has a much higher SME density than Poland, indicating probably the existence of many micro enterprises.

(TABLE 1)

Based on Table 1 and geographical and monetary criteria, we split our sample in three regional subsamples. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain and group 3 includes the Czech Republic and Poland. Table 2 summarizes the properties of our distress indicator for the overall sample and for the regional subsamples. As already mentioned, there is a bias due to the fact that in the beginning of the period (2000-2001), most firms in the database are survivors. It is immediately apparent that Eurozone distress rates are heightened in 2002-2003, are lower in 2004-2006 and are elevated again from 2007 onwards. This evidence is in accordance with the gloomy business climate in the early years of the last decade, which was followed by an impressive boom of the European economy in 2004-2006 and the subsequent slowdown that started in 2007.[13] The figures are somewhat different for group 3, which consists of two non-Eurozone members. This may be attributed to the fact that the credit supply by banks did not shrink in these countries in the years 2002-2003, as it did in most of the Eurozone countries. The distressed SMEs are 1.81% of all observations in the overall sample. Group 3 has the highest distress rate (2.4% of all firm-years).

(TABLE 2)

We should note that we follow Shumway (2001) and other authors and exclude financial firms from the sample (NACE[14] rev.2 codes from 64 to 68) due to the fact that financial firms have reporting practices that preclude combining them with other firms in models using financial information.

Because of the European focus of the study, we adopt the European Commission's definition for SMEs, instead of the more generic one of the Basel Committee, previously applied by Altman et al. (2010). We extract companies that meet the following requirements: (i) they have less than 250 employees and either annual turnover up to €50 million or total assets up to €43 million; (ii) there is no company with more than 25% participation in them; (iii) they do not have subsidiaries; (iv) they have up to ten shareholders; (v) they have at least two years of data available.

We need criteria (ii)-(iv) to ensure that the companies are independent.[15] Specifically, since we cannot track the subsidiaries and check if the companies still satisfy the SMEs criteria once we account for the subsidiaries' items, we need to exclude companies that have subsidiaries. Concerning criterion (iv), since the average number of shareholders in our sample is two, we exclude companies with more than ten shareholders as possible outliers. As to the last criterion, we keep companies with at least two years of data in order to be able to lag variables, calculate growth ratios and study the evolution of distress risk.

[13] The Eurozone insolvency index reported by Euler Hermes displays similar trends.

[14] NACE stands for "Nomenclature statistique des Nomenclature statistique des activités économiques dans la Communauté européenne".

[15] Altman et al. (2010) do not take into account the independence requirement when selecting their sample but try to control for it using a subsidiary dummy. They find that subsidiaries are less risky than non-subsidiaries. Small entities which are subsidiaries of large groups though can be very different from SMEs, especially when assessing their probability of distress. For example, Becchetti and Sierra (2003) find that group membership is inversely related to the probability of distress. Subsidiaries have access to financial, and other, resources of the group and can survive while experiencing poor financial performance. Moreover, the group may have reasons to support a subsidiary. Finally a subsidiary may be in distress as a result of group-wide distress.

After the initial extraction, we apply standard filtering and data cleaning techniques. We first check if missing values can be deduced from other items (i.e. if total assets are missing but fixed and current assets are available, we simply replace total assets with their sum). If the above method does not work, we exclude companies with missing values. We also exclude companies with errors in the data entered (i.e. companies that violate accounting identities).[16]

### C. Variables Selection

The factors that can lead SMEs to distress vary from firm-specific characteristics such as high debt to industry specific characteristics and macroeconomic effects such as high interest rates. To select among these factors, we take into account the models' stability, fit and parsimony as well as economic and statistical significance.

### C.1. Idiosyncratic Variables

Concerning the accounting data, we calculate financial ratios from nine categories: liquidity, profitability, interest coverage, leverage, activity, cash flow, growth (i.e. in sales or profits), asset utilization and employees efficiency.[17] We choose the ratios mainly based on economic intuition and suggestions from past literature. A list of the ratios examined is available upon request. As economic intuition suggests, we expect the probability of distress to be positively related to leverage and negatively related to all other ratio categories.

For the calculations, when denominators have zero values, we replace them with low values of €10 so that the ratios maintain their interpretation. Additionally, to ensure that statistical results are not heavily influenced by outliers, we set the bottom one percent to the first percentile and the top one percent to the

[16] Data availability constrains limit our initial dataset by around 25%.
[17] We do not examine ratios that have equity as one component because we characterize firms with negative equity that drop from the sample as distressed and in some cases such ratios have no economic meaning (i.e. equity to profits, when both equity and profits are negative). We need to note though that a distressed firm has negative equity in its latest balance sheet before leaving the sample, whereas in our models, we use accounting data lagged by one year.

ninety-ninth percentile.[18] Finally, as annual reports for SMEs become available with a significant time delay, we lag all ratios by 12 months in the estimations. This means that we assume that data for year $t-1$ become available at the end of year $t$.

After we calculate the candidate ratios, we follow a standard three-step procedure to select the best for our models. First, we follow Altman and Sabato's (2007) approach and find the AUC for each ratio, applying univariate analysis and keeping those with an AUC above 0.65. Second, we perform correlation analysis to avoid multicollinearity problems. When the correlation between two ratios is above 0.6, we keep the ratio with the highest AUC. If the difference in the AUC is small, we keep the ratio that was found to be important in previous studies. Finally, we apply a forward stepwise selection procedure of the remaining ratios.[19]

Table 3 reports summary statistics for the five ratios that are found to do the best overall job in predicting distress. A comparison of Panels B and C in Table 3 reveals the differences of distressed SMEs. Earnings before taxes to total assets differ substantially across the two groups suggesting the dominance of unprofitable SMEs in the distressed group. Another striking difference is that the distressed firm-years have on average around 130 times lower interest coverage compared to healthy firm-years. Short-term borrowing is also much higher in the case of distressed SMEs. Similarly, turnover to total liabilities ratio is around 180% higher in the healthy firm-years. Finally, the gap between distressed and healthy firm-years in the cash flow ratios indicates the importance of having high cash flows relative to current liabilities.

We do not expect large variations in the identified ratios when we repeat the selection exercise for different regions, since several past studies (for SMEs in the US, the UK and Sweden) also note their importance. We do expect differences in their coefficients though, since, when we look at Panels D, E and F, we notice differences in the ratios' sizes among regions.

---

[18] This popular technique is known as winsorizing and is widely used in the literature to avoid problems with outliers.

[19] Under this procedure, we start with no variables in the model, trying out the variables one by one and including them if they are statistically significant. We set the significance level at 10% and perform the likelihood ratio test which is more accurate than the standard Wald test.

(TABLE 3)

Recent studies though (Grunert et al., 2004; Altman et al., 2010) find that accounting ratios are not sufficient to predict SME distress risk and that including firm size and qualitative variables can improve predictive power. For this reason, we also account for size, industry type, number of shareholders, location, legal form and age.

The European Commission classifies SMEs into three groups based on their number of employees and turnover or total assets: medium-sized enterprises, small enterprises, and micro enterprises. As indicated in Panel A of Table 4, our sample is dominated by micro enterprises. In the sixth column of Panel A, the relationship between size and distress risk appears to be non-monotonic, with distress risk relatively stable for medium and small companies and higher for micro companies. This finding is consistent with other studies such as Dietsch and Petey (2004) and is also in line with the argument that smaller companies are more vulnerable to economic fluctuations. To test these predictions, we follow Altman et al. (2010) and employ the natural logarithm of total assets as a proxy for firm size. We also test for other specifications of size, such as the total turnover and the number of employees. Additionally, we examine interaction effects between size and the systematic variables that we introduce in the next subsection. For this purpose we use three size dummies (medium, small, micro) and combine them with the systematic variables to test the impact of macroeconomy on different size groups.

We also control for industry conditions using sector dummies to catch concentration effects. To construct our dummies, we use the NACE codes, which group industries into 21 major sectors. For estimation purposes though, this classification is too fine. The difficulty here relates to the grouping of sectors into wide sector classes in order to achieve an appropriate degree of homogeneity. It is true that such groupings can always be subject to a certain degree of arbitrariness. In our case, we follow an approach similar to Chava and Jarrow (2004) and form six wide sectors: (i) Sector 1: Agriculture, Mining and Manufacturing, (ii) Sector 2: Transportation, Communication and Utilities, (iii) Sector 3: Construction, (iv) Sector 4: Trade, (v) Sector 5: Accommodation and Food, and (vi) Sector 6: Other services. We select these wide sectors based on different regulatory environments, competition levels and product structures. We also test for alternative groupings but mostly get insignificant results for more detailed industry classifications.[20] Panel B of Table 4 shows the partitioning based on these wide sectors. This initial evidence shows that Accommodation and Food has the highest distress rate and Transportation, Communication and Utilities the lowest.

(TABLE 4)

Finally, we include a shareholders dummy (equal to one if the shareholders are more than two), a location dummy (equal to one if the SME is located in an urban area) and three legal form dummies in our models (for limited, unlimited and other legal forms). The average number of shareholders in our sample is two but 24% of SMEs have between three and ten shareholders. 14% of SMEs are located in big cities. 92% of SMEs have limited legal forms and few SMEs are cooperatives or partnerships. Generally, we expect SMEs with more shareholders to receive more capital injections in difficult times, thus have lower distress probabilities. Moreover, we expect SMEs in urban areas to be riskier due to higher competition among them. The intuition behind testing for the legal form of SMEs is that limited partners may be less interested to monitor firm performance compared to unlimited partners, leading limited SMEs more frequently to distress.[21] Whereas, as we show in the results part, we find support for our hypotheses concerning the number of shareholders and the location of SMEs, the coefficients of the legal dummies are statistically insignificant. Thus, we do not include them when reporting the results.

We lastly examine age for a smaller sample that we have the date of establishment available. Hudson (1987) finds that companies less than ten years old form most of distressed firms. In our sample, the average age at the time of distress is 11.9 years, whereas the average age for healthy firm-years is 15 years. Thus, we expect age to be negatively related to distress.

[20] I.e., we test for a more detailed classification of ten wide sectors, instead of six: 1. Agriculture; 2. Mining; 3. Manufacturing; 4. Utilities; 5. Construction; 6. Trade; 7. Accommodation and Food; 8. Transportation (and Storage); 9. Communication (and Information); 10. Other services. Our findings are not influenced and model performance remains the same.

[21] SMEs with unlimited partners can also proxy for family-owned firms, which are generally thought to be safer than other firms.

### C.2. Systematic Variables

In order to construct the systematic variables, we use data from Eurostat, the ECB, the World Bank and Datastream. Since these variables are often reported with a higher than annual frequency (quarterly, monthly or daily), we often need to annualize or take averages. We also usually lag them, in order to avoid causality considerations and because they are available for forecasting with a time delay. [22] So, we always use past realizations rather than expected values, assuming that these realizations are the best prediction we can have for the future. This is more appropriate for forecasting purposes since our objective is to predict distress at a certain point in time, given the definite information that we have available at this point, and because it is difficult to get reliable estimations for some systematic variables (i.e. FX rate or credit supply). These estimations usually differ among the various sources.

In our models, we use country-specific values and examine systematic variables from three categories: business cycle, credit conditions, and insolvency codes. In Appendix A we present the variables examined, their expected signs, calculation methods and number of lags, when applied.

Basing our predictions on economic rationale, we expect the probability of distress to be negatively related to business cycle variables such as appreciation of the local currency, disposable income or GDP growth and economic sentiment indicators. On the other hand, we expect it to be positively related to other business cycle variables such as country debt, inflation, oil price, unemployment and exchange rate volatility. European SMEs are mainly local market players and most often import raw materials and other supplies instead of exporting goods. Thus, an appreciation of the local currency makes these imports cheaper and lowers distress rates. Concerning disposable income, GDP, and economic sentiment, an increase in their values means better economic climate, thus it should be negatively related to distress. On the contrary, an increase in country debt, inflation, oil price, unemployment and exchange rate volatility signals uncertainly about future conditions of the economy and should be positively related to distress.

[22] We always test for different lags taking into account the economic rationale and the timing that the variables become available.

Concerning credit conditions, we expect the level of interest rates to be positively related to distress and bank lending to be negatively related to distress. An increase in interest makes it harder for SMEs to borrow whereas an increase in bank lending growth means better access to finance for them.

Finally, at this point, we need to elaborate on the effect of the bankruptcy laws on distress risk. Davydenko and Franks (2008) examine defaults in three European countries and find differences in insolvency codes among these countries to be important determinants of outcomes of the defaults. The World Bank measures the efficiency of insolvency codes in different countries based on the achieved recovery rate, which is the average percentage that claimants recover from an insolvent firm. The recovery rate depends on many factors such as the time it takes to resolve insolvency proceedings, costs and the outcome of the process. Generally, fast, low-cost proceedings and stronger creditor rights characterize the economies with high recovery rates. On the contrary, the more years to resolve an insolvency case, the less friendly the code is and the less likely for the firm to survive during the process. This is also obvious in Appendix B. Countries where the insolvency process takes longer, such as Czech Republic and Poland, score very low in the percentage of recovered amounts. The opposite is true for countries with fast proceedings such as UK and Germany. Thus, we expect distress rates to be negatively related to recovery rates and positively related to the time it takes to resolve insolvency proceedings. The above is also consistent with Acharya et al. (2011), who show that firms in countries with stronger creditor rights (thus higher recoveries) are more conservatively financed (i.e. have less debt).

In order to find among the systematic variables, the ones that significantly influence the probability of distress for SMEs, we follow a standard procedure. First, we fit the models using only accounting information.[23] Then, we run models that include the ratios and only one systematic variable at a time. We calculate the AUC for each of these models for the overall sample and for the

[23] In advanced top-down credit risk frameworks such as McKinsey and Co.'s CreditPortfolioView (Wilson 1997a and 1997b), macroeconomic factors are first fitted to aggregate distress rates and then the evolution of distress rates is simulated over time by generating macroeconomic shocks to the system. These simulated future distress rates, in turn, make it possible to obtain estimates of expected and unexpected losses for a credit portfolio, conditional on the current macroeconomic conditions. In this paper, our focus is the micro-level (individual SMEs). Finally, we acknowledge that there are feedback effects between the firm-specific factors and the macroeconomy but the modeling of these effects are beyond the scope of the study.

subsamples and keep the systematic variables that result to models with the highest AUCs. At this point, we need to account for correlation between the systematic variables. Correlations in this kind of variables are often high and lead to unreasonable signs of the estimated coefficients and to large changes in the values of these coefficients in response to small changes in the models' specifications. For this reason, between two systematic variables that have a correlation higher than 0.6, we keep the one that results in the model with the highest AUC.

When we fit our models using the regional subsamples, we anticipate systematic variables to vary across regions. Based on economic intuition, we suspect that group 2 (Italy, Portugal, Spain), which includes the peripheral economies of south Europe, is more exposed to the macroeconomy compared to group 1 (France, Germany, United Kingdom), that includes more stable economies. Also, we suspect that group 3 (Czech Republic, Poland) is exposed to currency risk since these countries are not members of the Eurozone.

Finally, we also form and examine interaction effects between industry dummies and systematic variables and firms' size and systematic variables. Generally, we predict that industries such as Construction and smaller SMEs are more vulnerable to the macroeconomy.

## III. The Results

In this part, we present empirical results and robustness checks. We first use the overall sample, then we split this sample in regional groups for the purpose of regional comparisons and finally, we split it in six rolling window periods to identify differences in coefficients and performance over time.

### A. Overall Sample

We estimate five models for the period 2000-2009. Model I includes only the idiosyncratic variables (accounting ratios, size, dummy for SMEs with more than two shareholders and dummy for SMEs in urban areas), model II includes both the idiosyncratic and the systematic variables, model III includes additionally the industry dummies, model IV includes some interaction terms, and, finally, model V includes age (available for a smaller sample). All models control for the duration effect, which is the "time at risk" of each firm in the sample.

### A.1. Empirical Results

Panel A of Table 5 presents the estimated coefficients and chi-square values for the five alternative model specifications. In model I, all firm-specific variables are significant and have the expected signs. Specifically, the probability of distress is negatively related to profitability (earnings before taxes to total assets), coverage (EBITDA to interest expenses), cash flow (cash flow to current liabilities) and activity (turnover to total liabilities) and positively related to leverage (current liabilities to total assets). Surprisingly, we do not find liquidity ratios as significant in the models. A logical explanation is that information contained in these ratios is proxied by others. I.e., the significance of current liabilities to total assets may indicate that SMEs rely more on short-term borrowing than cash holdings to finance their operational needs. The probability of distress is a decreasing function of the firm size (natural logarithm of total assets), indicating that as the firms become larger, they are less likely to undergo distress (see also Carling et al., 2007).[24] Two additional interesting findings in accordance with our predictions are that SMEs with less than three shareholders and SMEs in urban areas are riskier on average. The vast majority of SMEs have less than three shareholders (in our sample the percentage is 77%) but it seems that SMEs with more shareholders receive higher capital support in difficult times. This effect dominates the higher administrative costs that the existence of more shareholders may entail. A possible explanation for the higher risk of SMEs in urban areas is that these companies face higher competition (due to geographical proximity) and pay higher rents than their counterparts in the countryside. Another reason may be that owners of urban SMEs are less interested to support their enterprises in times of difficulties, since it is easier for them to shut down the business and find

[24] In unreported results, we also test for the nonlinear effects of size, by introducing the natural logarithm of squared total assets. We find a positive coefficient, indicating that for the very large SMEs, distress risk starts to increase, probably because these companies are more likely to be pursued in liquidation process by their creditors.

employment elsewhere. Thus, they may lead the firms to strategic distress. These effects seem to dominate the larger customer base available for urban SMEs.[25]

In model II, the firm-specific variables retain their significance and signs once the systematic variables are added. We identify five systematic variables as doing the best overall job in predicting distress, namely the FX rate change, the unemployment, the economic sentiment indicator, and the change in bank lending. All systematic variables have significant coefficients and the expected signs. As we hypothesized, an appreciation of the currency, an increase in the economic sentiment indicator and an increase in the lending by banks result in lower distress rates whereas an increase in unemployment and in the years to resolve insolvency result in higher distress rates. We should note though that we do not find the years to resolve insolvency to add predictive power in the regional models that we discuss later in the paper. This can partly be due to the fact that the regional groups are relatively homogeneous with respect to their insolvency regimes. To assess the usefulness of the systematic variables, we perform a likelihood ratio test for the nested models I and II. The null hypothesis that the coefficients of these variables are jointly equal to zero is strongly rejected, as indicated in Table 5.

Moving to model III, the firm-specific and systematic variables retain both their signs and significance and all industry dummies, except for industry 1 (Agriculture, Mining, Manufacturing)[26], enter with significant coefficients. Concerning the signs of the industry dummy coefficients, industries 2 (Utilities, Transportation, Communication) and 4 (Trade) are negatively related to distress and industry 3 (Construction) and 5 (Accommodation and Food) positively related to distress. To assess the usefulness of the industry dummies, we perform a likelihood ratio test for the nested models II and III. The null hypothesis is again rejected.

In model IV, we report results with interaction effects, in addition to the variables of model III. Specifically, we first test interaction effects between systematic variables and industry dummies, between

[25] Dietsch and Petey (2006) show something similar. Specifically, they find evidence from French SMEs that more attractive and wealthy regions demonstrate higher distress rates on average.

[26] We need to interpret this result with caution since the insignificant coefficient may result from the support packages provided under the European Union's Common Agricultural Policy during the period of the study.

systematic variables and size, and finally, between industry dummies and size. We find that the interaction effects that are most important in terms of performance improvement are between systematic variables and size dummies and report only these results for reasons of parsimony. From the coefficients of the interaction effects it is obvious that the distress probability of relatively larger SMEs (small and medium firms) is less sensitive to the systematic factors than the distress probability of the smallest SMEs (micro firms).[27] I.e. let us look how the effect of a bank lending change differs for the small and medium firms compared to micro firms. When we introduce interaction effects, the negative coefficient of the bank lending change increases in absolute size, demonstrating the increased sensitivity of micro firms to such a change. On the other hand, the additional effect of the bank lending change for small firms is positive (but still lower in absolute terms), and even more positive for medium firms. Thus, for the relatively larger SMEs, the same change in bank lending influences their distress probability less (but to the same direction) compared to micro firms. All other interaction effects display similar patterns with the exception of unemployment. Interestingly, the additional effect of unemployment for small and medium firms is of higher magnitude (-10.495 and -11.241 respectively) in absolute terms than the unemployment's coefficient (4.802). Thus, an increase in unemployment is positively related to the distress probability of micro firms but negatively related to the distress probability of small and medium firms. This may be due to the fact that in times of difficulty larger SMEs are more likely to fire employees in order to avoid bankruptcy and still be operational with fewer employees. Micro firms may not have such flexibility.

In model V, we introduce firm age and test its effect on distress probability for a slightly smaller sample for which we have available data on age. We find, in accordance with previous literature, that older firms are safer. Also, we follow Altman et al. (2010) and examine a non-monotonic effect of age. Specifically, Hudson (1987) finds evidence that start-ups are likely to have a "honeymoon" period of around three years before facing difficulties, provided they survive at their very beginning. To test this finding, Altman et al. (2010) introduce two dummy variables, one for firms from one to three years old

[27] Panel A of Table 4 provides details on the size classifications used here.

and one for firms between three and nine years old. They find a positive and statistically significant coefficient for the second dummy, exactly as in our case. We do not check the effect of the first dummy though because, in our case, we keep in the sample companies of at least two years due to the lags we apply.

At this point, we should clarify the role of the duration variable which controls for the survivorship bias described in Section I. We see that as the "time-at-risk", the duration in the sample increases, the distress probability increases as well. In our sample, the average age at the time of distress is 11.9 years, whereas in the overall sample the average age is 15 years. This further explains the positive sign of the duration variable. Since our sample period is 10 years, the maximum duration is 10, which is lower than the average age at distress.

Lastly, we notice that the pseudo-$R^2$ (McFadden's $R^2$) is increasing along the different model specifications, indicating a better fit as we add more variables. The pseudo-$R^2$ values may look low when compared to $R^2$ values of linear regression models, but such low values are normal in logistic regression (Hosmer and Lemeshow; 2000).

(TABLE 5)

### A.2. Robustness Checks

In order to evaluate the performance of our models, we perform in-sample and out-of-sample testing. We employ two widely used measures, the Hosmer and Lemeshow grouping based on estimated distress probabilities and the area under the Receiver Operating Characteristic (ROC) curve.

According to the Hosmer and Lemeshow method, the estimated distress probabilities for each year are ranked and divided into deciles. Out of the ten groups created (each one containing the 1/10 of the firms in that year), the first group has the smallest average estimated distress probability and the last the largest. Next, we aggregate the number of distressed firms in each decile for each year over 2000-2009 and calculate the corresponding percentages of the distressed firms in each decile.

The area under the ROC curve (AUC) is constructed from the estimated distress probabilities versus the actual status of the firms in each year for all possible cut-off probability values. Specifically, the curve plots the ratio of correctly classified distressed firms to actual distressed firms (sensitivity) and the ratio of wrongly classified healthy firms to actual healthy firms (1 - specificity) for all possible cut-offs. The AUC ranges from zero to one. A model with an AUC close to 0.5 is considered a random model with no discriminatory power. An AUC of 0.7 to 0.8 represents good discriminatory power, an AUC of 0.8 to 0.9 very good discriminatory power and an AUC over 0.9 is exceptional and extremely unusual. The AUC criterion is an improvement to the traditional classification tables that rely on a single cut-off point to classify distressed and healthy firms.[28] We should note at this point that the Hosmer and Lemeshow method assesses mainly calibration and the AUC assesses discrimination. We believe that our models' accuracy should be evaluated by considering both calibration and discrimination and for this reason we employ both tests.

Panel B of Table 5 presents the results of the in-sample tests. According to the Hosmer - Lemeshow grouping, the percentage of distressed firms in the last three deciles increases from model I to model II (75.83% to 76.59%). Also, the percentage of distressed firms in the first five deciles drops (11.38% to 11.09%). These show that adding the systematic variables improves performance both in terms of an increase in the correct classification of distressed firms and a decrease in the incorrect classification of healthy firms. AUC also increases from 0.8241 to 0.8382. This result is better than those achieved by previous studies in the literature. Specifically in Altman et al. (2010) this figure ranges between 0.78 and 0.80. When it comes to model III, it only modestly outperforms model II. Specifically, by taking industry effects into account, the AUC remains almost the same and the percentage of distressed firms in the last three deciles increases slightly (76.59% to 76.66%). Given these results, controlling for industry effects improves performance only marginally, once we have already accounted for systematic factors. When we add interaction effects between size and systematic

[28] Several statistics are equivalent to the AUC. The accuracy ratio (AR) can be derived from the AUC via a linear transformation (AR = 2AUC +1) and, thus, contains exactly the same information (Engelmann et al., 2003). The Gini coefficient, when defined with respect to the ROC curve, is identical in value to the AR, and, hence, also carries the same information. Finally, for continuous data, the AUC is equivalent to the Mann-Whitney U test (also known as Mann-Whitney-Wilcoxon or Wilcoxon rank-sum test).

factors, we notice a further increase in the percentage of distressed firms in the last three deciles (76.66% to 77.06%). AUC also increases from 0.8386 to 0.8431. Thus, once again, the involvement of systematic variables improves prediction accuracy. Moving to model V, it seems that age also helps slightly. We cannot though directly compare model IV to model V since model V is estimated with a smaller sample.

Panel C of Table 5 presents the results of the out-of-sample tests. Out-of-sample testing is challenging since improvements in the in-sample fit can be a result of over-fitting of the original data. We retain a random hold-out sample of 71,823 firms (304,037 firm-year observations), out of which 5,487 distressed, from the period 2000-2009 to perform out-of-sample validation. We use the coefficients estimates from the original models to predict distress for the hold-out sample and, as seen, all results follow the same patterns as for the in-sample tests.[29]

To further demonstrate the importance of systematic variables in distress prediction, Figure 1 plots the aggregate probability of distress for the whole sample period. We define the aggregate probability of distress as the simple average of the probabilities of distress of all firms in the sample each period. The shaded columns represent recession periods in the Eurozone as defined by OECD. The graph shows that in model I, where only firm-specific variables are included, the estimated probabilities of distress are relatively stable over time, following a smooth upward trend. It is the systematic variables (present in models II, III and IV) that shift the mean of the distress distribution and are able to capture distress clustering during recessions. When systematic variables are included, distress rates vary greatly with the business cycle, increasing on downturns and lowering on upturns. Once again, industry effects do not seem to provide additional improvements. These findings are in accordance with Jacobson et al. (2013) study for Sweden. They show that firm-specific variables account for the cross-section of the default

---

[29] The robustness checks provide evidence that the systematic variables and their interaction effects with size capture distress more successfully compared to the industry effects, which help only marginally. To test this finding, we run a model where we include only firm-specific information (model I) and the industry dummies. As expected, this model performs worse than model II, which includes firm-specific information and the systematic factors. Moreover, to exclude the possibility that the lower performance is due to inappropriate use of industry dummies, we use alternative industry classifications to construct our dummies and, still, get lower prediction power compared to model II. Finally, instead of negative equity, we use negative EBITDA to identify distressed firms. This alternative definition gives lower performance but same coefficient signs. The findings remain substantially similar under all tests and are available upon request.

distribution while the macroeconomic variables shift its mean in each period. Also, they find that industry effects offer small gains in their models in term of forecasting accuracy.

(FIGURE 1)

### B. Regional Subsamples

In addition to the overall sample, we estimate fitted models for the three subsamples presented in Section II. As explained, countries within the same group share common SME characteristics, thus, we examine each group separately.[30] We estimate two models for the period 2000-2009 for each group. Model I includes only the idiosyncratic variables, and model II includes both the idiosyncratic and the systematic variables. We do not report results with industry dummies, interaction variables and age for the sake of brevity. Results with these variables display similar patterns as those described above. Interestingly, we find that the firm-specific variables identified as the most important in predicting distress are exactly the same as for the overall sample. This is evidence that SMEs across Europe are sensitive to the same idiosyncratic factors. Concerning the systematic variables, as expected, there are regional variations in the vulnerabilities to systematic factors, indicating the need for region-specific models.

#### B.1. Empirical Results

Panel A of Table 6 presents the estimation results of the two models for group 1 (France, Germany, U.K.). The models are estimated from a sample of 165,786 SMEs (801,536 firm-year observations), which include 14,177 distressed SMEs. Again, all firm-specific variables are significant and have the expected signs. In group 1 model II, we find the bank lending and the GDP growth as the most useful macroeconomic variables in predicting distress. Both the bank lending and the GDP growth have significant coefficients and are, as expected, negatively related to the distress rate. Lower GDP growth

[30]We do not include country dummies in the regional models because such dummies do not have forecasting power. Instead, we estimate separate models for each country, but since findings are similar for countries of the same group, we present results of the regional models here.

means lower growth in sales of firms and thus an increased distress probability.

(TABLE 6)

Panel A of Table 7 presents the estimation results for group 2 (Italy, Portugal, Spain). The models are estimated from a sample of 429,978 SMEs (1,741,707 firm-year observations), which include 30,900 distressed SMEs. Again, the firm-specific variables are significant and have the expected signs, with small variations in the size of the coefficients from the coefficients of the overall models. In group 2 model II, we find four systematic factors as the most useful in predicting distress. The FX rate change, the bank lending change, and the economic sentiment indicator are negatively related to distress. On the other hand, the unemployment level is positively related to distress. It is interesting to note that, in accordance with our predictions, group 2 is vulnerable to more macroeconomic factors compared to group 1. The reason for this can be the generally worse economic climate in the economies of group 2 (Italy, Portugal, Spain) during the years of the study.

(TABLE 7)

Panel A of Table 8 presents the estimation results for group 3 (Czech Republic, Poland). The models are estimated from a sample of 48,470 SMEs (178,618 firm-year observations), which include 4,278 distressed SMEs. In group 3 model II, we find the FX volatility, the 10-year government bond yield and the GDP growth as the most useful systematic variables in predicting distress. With respect to the volatility of the exchange rate, higher volatility is positively related to distress (see also Nam et al., 2008). Interestingly, as we hypothesized, it seems that for the non- Eurozone countries of group 3, the stability of their national currencies plays a crucial role in the solvency of SMEs. This is presumably due to the fact that a very volatile FX rate in these economies increases instability, thus, uncertainty about future conditions of the economy. Concerning the 10-year government bond yield, it enters group 3 model II with a positive coefficient. Thus, a higher interest rate is positively related to distress. Government bond yields are systematically higher in the countries of group 3 compared to the rest of the sample for the years of the study, indicating the higher sovereign risk (country premium) for these economies. As before, GDP growth is negatively related to distress.

(TABLE 8)

### B.2. Robustness Checks

Panels B and C of Table 6 present the robustness exercises for group 1. The systematic factors improve performance slightly, since the percentage of distressed firms in the last three deciles increases from group 1 model I to group 1 model II (73.82% to 74.16%) and the AUC increases (0.8118 to 0.8254). Out-of-sample performance improvements are similar as in the case of in-sample results.

Panels B and C of Table 7 present the robustness exercises for group 2. Here, the inclusion of the systematic factors improves performance in terms of discriminatory power as the AUC increases from 0.8336 to 0.8482. Again out-of-sample results are similar.

Finally, Panels B and C of Table 8 present the results for group 3. According to the Hosmer-Lemeshow grouping, the percentage of distressed firms in the last three deciles increases from group 3 model I to group 3 model II by 2% (80.55% to 82.54%). Also, the percentage of distressed firms in the first five deciles drops (7.62% to 7.22%). AUC increases (0.8653 to 0.8749). Clearly, the systematic variables help in capturing distress risk compared to using only idiosyncratic information. The out-of-sample results give the same picture.

## C. Subperiods' Analysis

At this part, we estimate model III, which includes idiosyncratic and systematic variables as well as industry dummies, over different subperiods. Specifically, we estimate the model over four rolling windows, each five years long during the period 2002-2009. We perform this analysis for two reasons, first, in order to examine the stability of coefficients through time, and, secondly, to further test performance. We select not to examine a model that includes more variables for reasons of parsimony. This time, we evaluate predictive power over exactly the next year following each model's estimation period as well as over the last year of our sample (2009).

### C.1. Empirical Results

Panel A of Table 9 presents the estimation results of the four rolling windows over the period 2002-2009, as well as of the overall sample (model III) for comparison purposes. Coefficients of firm-specific variables are always significant and keep the same signs along the different windows, but there is relative variation in their magnitudes. The only puzzling result is the positive coefficient of size in the 2004-2008 window, but it seems that this result is sample specific. Coefficients of systematic variables follow the same patterns but display higher volatility, presumably as a result of the changing economic conditions during the period of the study. The years to resolve insolvency are negatively related to distress in the 2002-2006 window but this is probably also sample specific since distress rates are increasing quite impressively from 2002 to 2003 (Table 2) but insolvency regimes remain stable or improve. Finally, when it comes to industry dummies, their coefficients are often insignificant. Only industry 5 (Accommodation and Food) is always positively and significantly related to distress.

(TABLE 9)

### C.2. Robustness Checks

Panels B and C of Table 9 present the out-of-the-sample performance of the estimated rolling windows. Specifically, Panel B presents performance over the next year following the estimation period and Panel C presents performance over the last sample year (2009). In Panel A, the percentage of distressed SMEs in the last three deciles ranges from 72.93% - 78.15% and AUC ranges from 0.7825 – 0.8177. Similarly, in Panel B, the percentage of distressed SMEs in the last three deciles ranges from 71.93% - 72.93% and AUC ranges from 0.7795 – 0.7963.

## IV. Conclusions

The paper explores the performance of distress prediction hazard models for non-financial SMEs using a dataset from eight European countries over the ten-year period 2000-2009. The panel structure of the dataset allows us to exploit both the time-series and the cross-sectional dimension and differentiate between firm-specific, macroeconomic and industry effects.

We find that, in addition to financial indicators, whose importance has also been noted in past studies, SMEs in urban areas and SMEs with less than three shareholders have higher distress probabilities. We explore potential performance improvements reached by including systematic patterns and industry effects, in addition to firm-specific variables, to the distress prediction models and find that the exchange rate change, the economic sentiment, the bank lending conditions and the bankruptcy codes shift the mean of the distress distribution each period and are important determinants of the average distress rates in the economy. We validate the superiority of models that incorporate macroeconomic dependencies, suggested by previous research, also in the case of SMEs but do not find strong evidence that industry effects significantly improve prediction accuracy. We also examine interaction effects between SMEs' size and systematic variables and find that as SMEs become larger, they are less vulnerable to the macroeconomy. When we split our sample in regional groups, we identify regional variations in the importance of macro variables, but we show that SMEs across Europe are sensitive to the same firm-specific factors. Finally, we perform a rolling window analysis and find that whereas sensitivities to firm-specific factors remain relatively stable over time, coefficients of systematic variables are more volatile, reflecting changing macroeconomic conditions.

The paper's contribution to the field of distress prediction is multiple-fold. First, by using a dataset that includes a very high number of micro companies, we offer a better understanding of the European SME sector. Secondly, to our knowledge, we are the first to examine distress in a multi-country setting, allowing for cross-region comparisons. Finally, by considering systematic factors such as the macroeconomy, bank lending conditions and legal aspects, we uncover the main system-wide vulnerabilities of SMEs, both within Europe and among regions and capture distress comovements.

## References


Acharya, Viral V., Rangarajan K. Sundaram, and Kose John. (2011). "Cross-Country Variations in Capital Structure: The Role of Bankruptcy Costs." Journal of Financial Intermediation 20(1), 25-54.

Allen, Linda, and Anthony Saunders. (2004). "Incorporating Systemic Influences Into Risk Measurements: A Survey of the Literature." Journal of Financial Services Research 26(2), 161-192.

Altman, Edward I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." The Journal of Finance 23(4), 589-609.

Altman, Edward I., and Edith Hotchkiss. (2005). Corporate Financial Distress and Bankruptcy, 3rd ed. Wiley.

Altman, Edward I., and Paul Narayanan. (1997). "An International Survey of Business Failure Classification Models." Financial Markets, Institutions and Instruments 6(2), 1-57.

Altman, Edward I., and Gabriele Sabato. (2007). "Modelling Credit Risk for SMEs: Evidence from the U.S. Market." Abacus 43(3), 332-357.

Altman, Edward I., Gabriele Sabato, and Nicholas Wilson. (2010). "The Value of Non-Financial Information in SME Risk Management." Journal of Credit Risk 6(2), 95-127.

Altman, Edward I., and Anthony Saunders. (1998). "Credit Risk Measurement: Developments Over the last 20 years." Journal of Banking and Finance 21(11). 1721–1742.

Basel Committee on Banking Supervision. (2006). "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version". Basel, Switzerland, Bank for International Settlements.

Becchetti, Leonardo, and Jaime Sierra. (2003). "Bankruptcy risk and productive efficiency in manufacturing firms." Journal of Banking and Finance 27(11), 2099–2120.

Beck, Thorsten, Asli Demirguc-Kunt, and Vojislav Maksimovic. (2008). "Financing Patterns around the World: Are Small Firms Different?" Journal of Financial Economics 89(3), 467-487.

Berger, Allen N., and Gregory F. Udell. (1998). "The Economics of Small Business Finance: The Roles of Private Equity and Debt Markets in the Financial Growth Cycle." Journal of Banking and Finance 22(6-8), 613-673.

Berger, Allen N., and Gregory F. Udell. (2006). "A More Complete Conceptual Framework about SME Finance." Journal of Banking and Finance 30(11), 2945-2966.

Bruneau, Catherine, Olivier de Bandt, and Widad El Amri. (2012). "Macroeconomic Fluctuations and Corporate Financial Fragility." Journal of Financial Stability 8(4), 219-235.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. (2008). "In Search of Distress Risk." The Journal of Finance 63(6), 2899-2939.

Carling, Kenneth, Tor Jacobson, Jesper Lindé, and Kasper Roszbach. (2007). "Corporate Credit Risk Modeling and the Macroeconomy." Journal of Banking and Finance 31(3), 845–868.

Chava, Sudheer, and Robert Jarrow. (2004). "Bankruptcy Prediction with Industry Effects." Review of Finance 8(4), 537–569.

Cox, David R. (1972). "Regression Models and Life-Tables." Journal of the Royal Statistical Society Series B 34(2), 187-220.

Davydenko, Sergei A. (2012). "When Do Firms Default? A Study of the Default Boundary". Working Paper.

Davydenko, Sergei A., and Julian R. Franks. (2008). "Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the U.K." The Journal of Finance 63(2), 568-608.

DeYoung, Robert, Dennis Glennon, and Peter Nigro. (2008). "Borrower–Lender Distance, Credit Scoring, and Loan Performance: Evidence from informational-opaque small business borrowers." Journal of Financial Intermediation 17(1), 113–143.

Dietsch, Michel, and Joël Petey. (2004). "Should SME Exposures be treated as Retail or Corporate Exposures? A Comparative Analysis of Default Probabilities and Asset Correlations in French and German SMEs." Journal of Banking and Finance 28(4), 773-788.

Duffie, Darrel, Leandro Saita, and Ke Wang. (2007). "Multi-period Corporate Default Prediction with Stochastic Covariates." Journal of Financial Economics 83(3), 635–665.

Edminster, Robert O. (1972). "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction." The Journal of Financial and Quantitative Analysis 7(2), 1477-1493.

Engelmann, Bernd, Evelyn Hayden, and Dirk Tasche. (2003). "Testing Rating Accuracy." Journal of Risk 16(1), 82-86.

Euler Hermes (2011). "Macroeconomic Risk and Insolvency Outlook." Economic Outlook no. 9/2011.

European Commission (2007). "Observatory of European SMEs." Analytical Report, Flash Eurobarometer 196.

Froot, Kenneth A. (1989). "Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data." Journal of Financial and Quantitative Analysis 24(3), 333-355.

Gilson, Stuart C., and Michael R. Vetsuypens. (1993). "CEO Compensation in Financially Distressed Firms: An Empirical Analysis." The Journal of Finance 48(2), 425-458.

Glennon, Dennis, and Peter Nigro. (2005). "Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach." Journal of Money, Credit and Banking 37(5), 923–947.

Grunert Jens, Lars Norden, and Martin Weber. (2005). "The role of non-financial factors in internal credit ratings." Journal of Banking and Finance 29(2), 509-531.

Headd, Brian. (2003). "Redefining Business Success: Distinguishing Between Closure and Failure." Small Business Economics 21(1), 51–61.

Hillegeist, Stephen A., Elizabeth K. Keating, Donald P. Cram, and Kyle G. Lundstedt. (2004). "Assessing the Probability of Bankruptcy." Review of Accounting Studies 9(1), 5-34.

Hosmer, David W., and Stanley Lemeshow. (2000). Applied Logistic Regression, 2nd ed. Wiley.

Jacobson, Tor, Jesper Lindé, and Kasper Roszbach. (2005). “Exploring Interactions between Real Activity and the Financial Stance.” Journal of Financial Stability 1(3), 308-341.

Jacobson, Tor, Jesper Lindé, and Kasper Roszbach. (2013). “Firm Default and Aggregate Fluctuations.” Journal of the European Economic Association 11(4), 942-972.

Koopman, Siem Jan, Roman Kräussl, André Lucas, and André B. Monteiro. (2009). “Credit Cycles and Macrofundamentals.” Journal of Empirical Finance 16(1), 42-54.

Merton, Robert C. (1974). “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.” The Journal of Finance 29(2), 449–470.

Nam, Chae Woo, Tong Suk Kim, Nam Jung Park, and Hoe Kyung Lee. (2008). “Bankruptcy Prediction Using a Discrete-Time Duration Model Incorporating Temporal and Macroeconomic Dependencies.” Journal of Forecasting 27(6), 493–506.

OECD (2005). OECD SME and Entrepreneurship Outlook 2005.

Ross, Stephen A., Randolph Westerfield, and Jeffrey Jaffe. (2010). Corporate Finance, 9th ed. McGraw-Hill/Irwin Series.

Shumway, Tyler. (2001). “Forecasting Bankruptcy More Accurately: A Simple Hazard Model.” Journal of Business 74(1), 101-124.

Stein, Jeremy C. (2002). “Information Production and Capital Allocation: Decentralized versus Hierarchical Firms.” The Journal of Finance 57(5), 1891-1921.

Watson, John, Jim Everett. (1993). “Defining Small Business Failure.” International Small Business Journal 3(11), 35–48.

White, Halbert. (1994). Estimation, Inference and Specification Analysis. Cambridge University Press.

Wilson, Thomas C. (1997a). “Portfolio Credit Risk (I).” Risk 10(9), 111-117.

Wilson, Thomas C. (1997b). “Portfolio Credit Risk (II).” Risk 10(10), 56-61.

Wooldridge, Jeffrey M. (2002). Econometric Analysis of Cross-section and Panel Data. MIT press.

World Bank. (2013). “Doing Business, Smarter Regulations for Small and Medium-Sized Enterprises.” Report 2013.

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

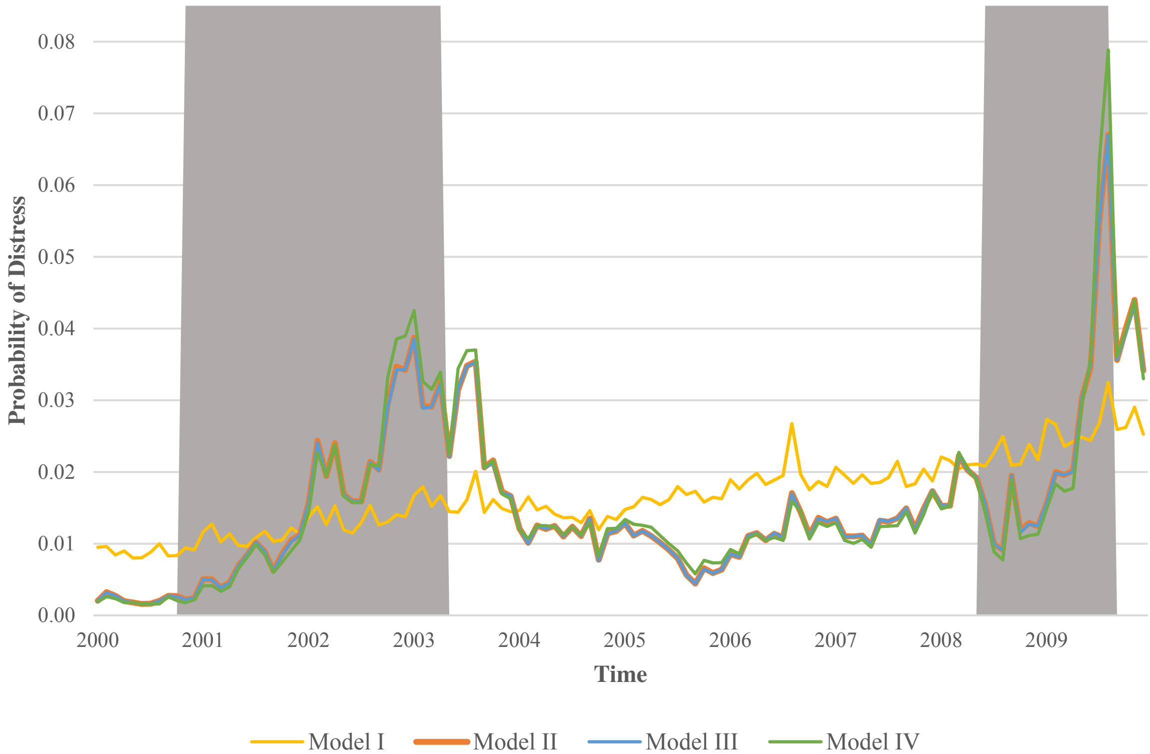


**Figure 1. Aggregate Probability of Distress.** The figure plots the aggregate probability of distress for firms in the overall sample. We define it as the simple average probability of distress of all firms. The columns denote recession periods in the Euro area, as indicated by OECD.

**Table 1**

**Key Indicators**

The role of SMEs varies substantially across the EU. The table gives an overview of SMEs in the EU27 and in the countries of our specific interest. The first column gives the contribution of SMEs to employment, the second the contribution to the value-added in the economy and the third the density of SMEs per 1,000 inhabitants.

| | (%) of employment | (%) of value added | Number per 1000 inhabitants |
|---|---|---|---|
| EU27 | 67.1 | 57.6 | 39.9 |
| Italy | 81.3 | 70.9 | 65.3 |
| Portugal | 82.0 | 67.8 | 80.5 |
| Spain | 78.7 | 68.5 | 59.1 |
| France | 61.4 | 54.2 | 36.3 |
| Germany | 60.6 | 53.2 | 20.0 |
| United Kingdom | 54.0 | 51.0 | 25.6 |
| Czech Republic | 68.9 | 56.7 | 86.0 |
| Poland | 69.8 | 48.4 | 36.8 |

**Table 2**
**Distressed SMEs as Percentage of Total SMEs**

The table summarizes the properties of our distress indicator for the overall sample and for the regional subsamples. It gives the number of total SMEs in the beginning of the year, the number of distressed SMEs during the year and the distress rate per year.

| | Overall | | | Group 1 | | | Group 2 | | | Group 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Total | Distressed | (%) | Total | Distressed | (%) | Total | Distressed | (%) | Total | Distressed | (%) |
| 2000 | 149,023 | 0 | 0.00 | 82,666 | 0 | 0.00 | 65,576 | 0 | 0.00 | 781 | 0 | 0.00% |
| 2001 | 176,351 | 192 | 0.11 | 92,348 | 185 | 0.20 | 81,782 | 6 | 0.01 | 2,221 | 1 | 0.05% |
| 2002 | 204,531 | 3,802 | 1.86 | 99,815 | 2,125 | 2.13 | 100,466 | 1,649 | 1.64 | 4,250 | 28 | 0.66% |
| 2003 | 194,768 | 5,961 | 3.06 | 91,761 | 4,003 | 4.36 | 94,857 | 1,935 | 2.04 | 8,150 | 23 | 0.28% |
| 2004 | 146,877 | 1,250 | 0.85 | 52,031 | 865 | 1.66 | 81,727 | 331 | 0.41 | 13,119 | 54 | 0.41% |
| 2005 | 167,837 | 1,403 | 0.84 | 53,609 | 822 | 1.53 | 99,053 | 377 | 0.38 | 15,175 | 204 | 1.34% |
| 2006 | 256,732 | 1,873 | 0.73 | 70,242 | 902 | 1.28 | 164,105 | 734 | 0.45 | 22,385 | 237 | 1.06% |
| 2007 | 463,732 | 8,134 | 1.75 | 95,393 | 1,600 | 1.68 | 331,731 | 5,932 | 1.79 | 36,608 | 602 | 1.64% |
| 2008 | 498,358 | 9,194 | 1.84 | 88,606 | 1,427 | 1.61 | 369,487 | 6,977 | 1.89 | 40,265 | 790 | 1.96% |
| 2009 | 463,652 | 17,546 | 3.78 | 75,065 | 2,248 | 2.99 | 352,923 | 12,959 | 3.67 | 35,664 | 2,339 | 6.56% |
| Obser. | 2,721,861 | 49,355 | 1.81 | 801,536 | 14,177 | 1.77 | 1,741,707 | 30,900 | 1.77 | 178,618 | 4,278 | 2.40% |

**Table 3**
**Summary Statistics**

The table reports summary statistics for all of the accounting ratios used to forecast distress. Each observation represents a particular firm in a particular year. Panel A describes the distributions of the ratios in all firm-years, Panel B describes the sample of healthy years, and Panel C describes the distressed years. Panels D, E and F describe the distributions for Groups 1, 2 and 3 respectively. The sample period is from 2000 to 2009. All ratios are truncated at the ninety-ninth and first percentiles.

| | Earnings before taxes to total assets | EBITDA to interest expenses | Current liabilities to total assets | Cash flow to current liabilities | Turnover to total liabilities |
|---|---|---|---|---|---|
| | | | Panel A. Entire data set | | |
| Mean | 0.05 | 687.28 | 0.61 | 0.31 | 3.60 |
| Median | 0.04 | 7.00 | 0.59 | 0.12 | 2.57 |
| Std.Dev. | 0.17 | 2,927.14 | 0.34 | 0.86 | 4.13 |
| Min | -0.85 | -2,600.00 | 0.00 | -1.17 | 0.09 |
| Max | 0.63 | 21,200.00 | 2.27 | 7.00 | 30.59 |
| N: 2,721,861 | | | | | |
| | | | Panel B. Healthy Group | | |
| Mean | 0.05 | 699.87 | 0.60 | 0.31 | 3.63 |
| Median | 0.04 | 7.29 | 0.59 | 0.13 | 2.59 |
| Std.Dev. | 0.17 | 2,945.99 | 0.33 | 0.86 | 4.15 |
| N: 2,672,506 | | | | | |
| | | | Panel C. Distressed Group | | |
| Mean | -0.13 | 5.39 | 1.02 | -0.01 | 2.04 |
| Median | -0.04 | 0.65 | 0.92 | 0.00 | 1.42 |
| Std.Dev. | 0.29 | 1,448.37 | 0.56 | 0.59 | 2.50 |
| N: 49,355 | | | | | |
| | | | Panel D. Group 1 | | |
| Mean | 0.08 | 1,064.80 | 0.61 | 0.32 | 3.76 |
| Median | 0.06 | 12.75 | 0.60 | 0.16 | 3.18 |
| Std.Dev. | 0.17 | 3,682.35 | 0.29 | 0.79 | 2.86 |
| N: 801,536 | | | | | |
| | | | Panel E. Group 2 | | |
| Mean | 0.03 | 493.67 | 0.61 | 0.28 | 3.25 |
| Median | 0.03 | 5.18 | 0.60 | 0.10 | 2.10 |
| Std.Dev. | 0.17 | 2,426.19 | 0.35 | 0.85 | 4.22 |
| N: 1,741,707 | | | | | |
| | | | Panel F. Group 3 | | |
| Mean | 0.09 | 881.04 | 0.58 | 0.55 | 6.32 |
| Median | 0.07 | 13.00 | 0.53 | 0.20 | 4.31 |
| Std.Dev. | 0.23 | 3,357.95 | 0.41 | 1.19 | 6.39 |
| N: 178,618 | | | | | |

**Table 4**

**SMEs by Size and Industry**

Panel A. Size classification

The panel shows the classification of SMEs by size. The first column shows the size classes. The second column shows the firms available in each class, the third column shows the percentage of firms available in each class, the fourth column shows the number of firm-years available in each class and the fifth column shows the distressed firm-years available in each class. Finally the sixth column shows the distress rate as a percentage of total firm-years in each class.

| Size | | | | | Firms | (%) firms | Firm-years | Distressed | (%) distressed |
|---|---|---|---|---|---|---|---|---|---|
| Cat. | Employees | Turnover | or | Assets | | | | | |
| Medium | < 250 | ≤ € 50 m | | ≤ € 43 m | 21,408 | 3.32 | 123,123 | 1,815 | 1.47 |
| Small | < 50 | ≤ € 10 m | | ≤ € 10 m | 167,381 | 25.98 | 906,392 | 13,183 | 1.45 |
| Micro | < 10 | ≤ € 2 m | | ≤ € 2 m | 455,445 | 70.70 | 1,692,346 | 34,357 | 2.03 |
| Total | | | | | 644,234 | 100.00 | 2,721,861 | 49,355 | 1.81 |

Panel B. Industry classification (wide sectors)

The panel shows the classification of SMEs by wide industry sectors. The first column shows the sectors. The second column shows the firms available in each sector, the third column shows the percentage of firms available in each sector, the fourth column shows the number of firm-years available in each sector and the fifth column shows the distressed firm-years available in each sector. Finally the sixth column shows the distress rate as a percentage of total firm-years in each

| Sector | Firms | (%) firms | Firm-years | Distressed | (%) distressed |
|---|---|---|---|---|---|
| 1. Agriculture, Mining and Manufacturing | 133,746 | 20.76 | 608,696 | 9,815 | 1.61 |
| 2. Transportation, Communication and Utilities | 45,413 | 7.05 | 182,180 | 2,827 | 1.55 |
| 3. Construction | 113,147 | 17.56 | 482,031 | 9,170 | 1.90 |
| 4. Trade | 214,061 | 33.23 | 946,368 | 16,291 | 1.72 |
| 5. Accommodation and Food | 36,235 | 5.62 | 128,225 | 3,691 | 2.88 |
| 6. Other services | 101,632 | 15.78 | 374,361 | 7,561 | 2.02 |
| Total | 644,234 | 100.00 | 2,721,861 | 49,355 | 1.81 |

## Table 5

**Overall Sample (8 countries)**

Panel A. Estimation results

The models are estimated for 2000-2009 data with yearly observations using the multi-period logit technique. All firm-specific variables are lagged by one year. The data set includes SMEs from eight European economies. Parameter estimates are given first followed by chi-square values in parentheses.

| | Model I | | Model II | | Model III | | Model IV | | Model V | |
|---|---|---|---|---|---|---|---|---|---|---|
| Earnings before taxes to total assets | -0.755*** | (-15.65) | -0.770*** | (-15.89) | -0.763*** | (-15.53) | -0.779*** | (-15.99) | -0.777*** | (-14.92) |
| EBITDA to interest expenses | -0.0000453*** | (-14.98) | -0.0000450*** | (-14.49) | -0.0000451*** | (-14.58) | -0.0000441*** | (-14.38) | -0.000045*** | (-14.20) |
| Current liabilities to total assets | 1.381*** | (101.27) | 1.420*** | (103.81) | 1.417*** | (102.97) | 1.409*** | (101.04) | 1.38*** | (95.92) |
| Cash flow to current liabilities | -0.480*** | (-8.99) | -0.485*** | (-9.14) | -0.491*** | (-9.16) | -0.475*** | (-9.00) | -0.517*** | (-9.22) |
| Turnover to total liabilities | -0.182*** | (-36.96) | -0.177*** | (-36.08) | -0.176*** | (-35.56) | -0.182*** | (-35.56) | -0.187*** | (-35.64) |
| Size (ln(totals assets)) | -0.127*** | (-30.88) | -0.0940*** | (-22.95) | -0.0913*** | (-22.14) | -0.109*** | (-23.13) | -0.097*** | (-20.18) |
| Dummy equal to 1 if shareholders are more than 2 | -0.291*** | (-23.53) | -0.274*** | (-21.99) | -0.272*** | (-21.76) | -0.270*** | (-21.50) | -0.225*** | (-17.54) |
| Dummy equal to 1 if SME is located in an urban area | 0.132*** | (10.24) | 0.141*** | (10.85) | 0.144*** | (11.01) | 0.153*** | (11.54) | 0.175*** | (13.01) |
| Duration | 0.264*** | (145.44) | 0.227*** | (108.52) | 0.228*** | (108.86) | 0.229*** | (107.87) | 0.284*** | (117.75) |
| FX rate (% change) | | | -1,686.8*** | (-59.04) | -1,689.9*** | (-59.01) | -2,627.20*** | (-68.63) | -2,695.82*** | (-69.51) |
| Unemployment | | | 1.883*** | (12.39) | 1.914*** | (12.58) | 4.802*** | (28.84) | 4.345*** | (25.82) |
| Economic sentiment indicator | | | -0.0259*** | (-35.03) | -0.0258*** | (-34.90) | -0.0388*** | (-48.72) | -0.0386*** | (-46.76) |
| Loans granted to non-financial sector (% change) | | | -4.414*** | (-58.29) | -4.407*** | (-58.07) | -5.246*** | (-53.94) | -5.226*** | (-54.84) |
| Years to resolve insolvency proceedings | | | 0.0949*** | (27.40) | 0.0958*** | (27.57) | 0.1211*** | (25.80) | 0.1209*** | (25.75) |
| Industry 1 (Agriculture, Mining, Manufacturing) | | | | | | | 0.0442*** | (3.48) | 0.0938*** | (7.26) |
| Industry 2 (Utilities, Transportation, Communication) | | | | | -0.0762*** | (-3.56) | | | | |
| Industry 3 (Construction) | | | | | 0.0798*** | (5.84) | 0.1035*** | (8.06) | 0.0782*** | (6.01) |
| Industry 4 (Trade) | | | | | -0.0295* | (-2.50) | | | | |
| Industry 5 (Accommodation and Food) | | | | | 0.212*** | (10.18) | 0.251*** | (12.25) | 0.3169*** | (15.49) |
| Small firm* FX rate (% change) | | | | | | | 1,796.63*** | (35.16) | 1,737.00*** | (33.17) |
| Small firm* unemployment | | | | | | | -10.495*** | (-37.76) | -10.591*** | (-37.41) |
| Small firm* economic sentiment indicator | | | | | | | 0.0146*** | (30.15) | 0.0151*** | (30.47) |
| Small firm* loans to non-financial sector (% ch.) | | | | | | | 1.771*** | (10.93) | 1.639*** | (10.24) |
| Small firm* years to resolve insolvency proceedings | | | | | | | -0.0493*** | (-6.79) | -0.0673*** | (-8.95) |
| Medium firm* FX rate (% change) | | | | | | | 1,936.71*** | (20.51) | 1,975.46*** | (19.98) |
| Medium firm* unemployment | | | | | | | -11.241*** | (-15.40) | -12.091*** | (-15.97) |
| Medium firm* economic sentiment indicator | | | | | | | 0.0174*** | (16.96) | 0.0196*** | (18.36) |
| Medium firm* loans to non-financial sector (% ch.) | | | | | | | 4.084*** | (14.02) | 3.766*** | (12.94) |
| Medium firm* years to resolve insolvency proceedings | | | | | | | -0.1392*** | (-9.81) | -0.1605*** | (-10.94) |
| Age | | | | | | | | | -0.0133*** | (-17.30) |
| Age (3-9) | | | | | | | | | 0.5501*** | (43.76) |
| Constant | -4.675*** | (-140.65) | -2.479*** | (-29.79) | -2.523*** | (-30.26) | -1.6732*** | (-19.07) | -1.9558*** | (-21.45) |

| | Model I | Model II | Model III | Model IV | Model V |
|---|---|---|---|---|---|
| Firm-year observations | 2,721,861 | 2,721,861 | 2,721,861 | 2,721,861 | 2,652,157 |
| Firms | 644,234 | 644,234 | 644,234 | 644,234 | 620,872 |
| Distressed firms | 49,355 | 49,355 | 49,355 | 49,355 | 47,841 |
| Pseudo R-squared | 0.147 | 0.171 | 0.171 | 0.178 | 0.187 |
| Log likelihood | -210,601.30 | -204,638.50 | -204,538.30 | 202,880.11 | 194,837.44 |
| Wald test | 78,110.8*** | 84,259.5*** | 84,526.8*** | 85,305.9 | 81,789.3*** |
| Likelihood ratio test | | 11,925.57*** | 200.45*** | 3,316.36 | 16,085.34*** |
| * p<0.05, ** p<0.01, *** p<0.001 | | | | | |
| Panel B. In-sample prediction tests | | | | | |
| Hosmer-Lemeshow test: Percentage of the 49,355 distressed firms predicted in each decile in the year of their distress. | | | | | |
| Decile | | | | | |
| 1 to 5 | 11.38% | 11.09% | 10.96% | 10.67% | 10.24% |
| 6 | 5.28% | 5.16% | 5.09% | 5.01% | 5.27% |
| 7 | 7.50% | 7.15% | 7.28% | 7.27% | 7.33% |
| 8 | 11.20% | 11.16% | 11.25% | 10.91% | 10.33% |
| 9 | 17.46% | 17.86% | 17.84% | 17.83% | 17.34% |
| 10 | 47.17% | 47.58% | 47.57% | 48.32% | 49.49% |
| 8 to 10 | 75.83% | 76.59% | 76.66% | 77.06% | 77.16% |
| 9 to 10 | 64.63% | 65.44% | 65.41% | 66.15% | 66.83% |
| Area under the ROC curve | 0.8241 | 0.8382 | 0.8386 | 0.8431 | 0.8571 |
| Panel C. Out-of-sample prediction tests | | | | | |
| A hold-out sample of 71,823 European SMEs (304,037 firm-year observations) is used. | | | | | |
| Hosmer-Lemeshow test: Percentage of the 5,487 distressed firms predicted in each decile in the year of their distress. | | | | | |
| Decile | | | | | |
| 1 to 5 | 11.46% | 11.35% | 11.26% | 10.30% | 10.15% |
| 6 | 5.41% | 4.76% | 4.59% | 5.41% | 5.39% |
| 7 | 7.58% | 7.62% | 7.91% | 6.74% | 7.63% |
| 8 | 11.24% | 10.53% | 10.66% | 11.45% | 10.71% |
| 9 | 17.51% | 18.70% | 18.53% | 18.15% | 17.16% |
| 10 | 46.78% | 47.04% | 47.04% | 47.95% | 48.95% |
| 8 to 10 | 75.54% | 76.27% | 76.23% | 77.55% | 76.82% |
| 9 to 10 | 64.30% | 65.74% | 65.57% | 66.10% | 66.11% |
| Area under the ROC curve | 0.8234 | 0.8369 | 0.8373 | 0.8437 | 0.8472 |

# Table 6

## Group 1 (France, Germany, United Kingdom)

Panel A: Estimation results

The models are estimated for 2000-2009 data with yearly observations using the multi-period logit technique. All firm-specific va are lagged by one year. The data set is limited to non-financial French, German and British SMEs. Parameter estimates are give followed by chi-square values in parentheses. There are 165,786 firms in the sample (801,536 firm-year observations) out of 14,177 distressed.

| | Group 1 Model I | | Group 1 Mode | |
|---|---|---|---|---|
| Earnings before taxes to total assets | -1.081*** | (-15.45) | -1.077*** | (- |
| EBITDA to interest expenses | -0.0000482*** | (-11.49) | -0.0000483*** | (- |
| Current liabilities to total assets | 1.863*** | (70.42) | 1.916*** | (7 |
| Cash flow to current liabilities | -0.173** | (-2.69) | -0.196** | ( |
| Turnover to total liabilities | -0.104*** | (-14.23) | -0.101*** | (- |
| Size (ln(totals assets)) | -0.0345*** | (-4.54) | -0.00559 | ( |
| Dummy equal to 1 if shareholders are more than 2 | -0.0880*** | (-3.94) | -0.0812*** | ( |
| Dummy equal to 1 if SME is located in an urban area | 0.184*** | (5.96) | 0.174*** | |
| Duration | 0.189*** | (59.93) | 0.168*** | (4 |
| Loans granted to non-financial sector (% change) | | | -4.611*** | (-2 |
| GDP growth (% change) | | | -5.595*** | ( |
| Industry 1 (Agriculture, Mining, Manufacturing) | | | | |
| Industry 2 (Transportation, Communication, Utilities) | | | | |
| Industry 3 (Construction) | | | | |
| Industry 5 (Accommodation and Food) | | | | |
| Constant | -5.499*** | (-82.31) | -5.293*** | (-8 |
| Firm-year observations | 801,536 | | 801,536 | |
| Firms | 165,786 | | 165,786 | |
| Distressed firms | 14,177 | | 14,177 | |
| Pseudo R-squared | 0.136 | | 0.150 | |
| Log likelihood | -61,573.5 | | -60,538.7 | |
| Wald test | 17,957.5*** | | 20,225.9*** | |
| Likelihood ratio test | | | 2,069.68*** | |

* p<0.05, ** p<0.01, *** p<0.001

Panel B. In-sample prediction tests

Hosmer-Lemeshow test: Percentage of the 14,177 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 13.74% | 13.85% |
| 8 | 9.46% | 9.47% |
| 9 | 13.80% | 14.02% |
| 10 | 50.55% | 50.67% |
| 8 to 10 | 73.82% | 74.16% |
| Area under the ROC curve | 0.8118 | 0.8254 |

Panel C: Out-of-sample prediction tests

A hold-out sample of 18,449 French, German and British SMEs (88,957 firm-year observations) is used.

Hosmer-Lemeshow test: Percentage of the 1,626 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 13.59% | 13.47% |
| 8 | 9.41% | 8.79% |
| 9 | 14.21% | 15.01% |
| 10 | 48.89% | 48.95% |
| 8 to 10 | 72.51% | 72.76% |
| Area under the ROC curve | 0.8056 | 0.8236 |

## Table 7

### Group 2 (Italy, Portugal, Spain)

Panel A. Estimation results

The models are estimated for 2000-2009 data with yearly observations using the multi-period logit technique. All firm-specific variables are lagged by one year. The data set is limited to non-financial Italian, Portuguese and Spanish SMEs. Parameter estimates are given first followed by chi-square values in parentheses. There are 429,978 firms in the sample (1,741,707 firm-year observations) out of which 30,900 distressed.

| | Group 2 Model I | | Group 2 Model II | |
|---|---|---|---|---|
| Earnings before taxes to total assets | -0.704*** | (-11.06) | -0.677*** | (-10.61) |
| EBITDA to interest expenses | -0.0000485*** | (-9.87) | -0.0000468*** | (-9.82) |
| Current liabilities to total assets | 1.162*** | (66.78) | 1.217*** | (69.66) |
| Cash flow to current liabilities | -0.620*** | (-8.96) | -0.650*** | (-9.43) |
| Turnover to total liabilities | -0.272*** | (-32.51) | -0.249*** | (-30.41) |
| Size (ln(totals assets)) | -0.180*** | (-32.85) | -0.107*** | (-17.72) |
| Dummy equal to 1 if shareholders are more than 2 | -0.387*** | (-24.19) | -0.324*** | (-19.96) |
| Dummy equal to 1 if SME is located in an urban area | 0.0776*** | (4.94) | 0.101*** | (6.43) |
| Duration | 0.305*** | (128.02) | 0.206*** | (67.49) |
| FX rate (% change) | | | -2276.6*** | (-44.99) |
| Unemployment | | | 6.176*** | (24.91) |
| Loans granted to non-financial sector (% change) | | | -3.378*** | (-30.14) |
| Economic sentiment | | | -0.0256*** | (-21.08) |
| Industry 1 (Agriculture, Mining, Manufacturing) | | | | |
| Industry 3 (Construction) | | | | |
| Industry 4 (Trade) | | | | |
| Industry 5 (Accommodation and Food) | | | | |
| Constant | -4.230*** | (-99.51) | -2.921*** | (-19.57) |
| Firm-year observations | 1,741,707 | | 1,741,707 | |
| Firms | 429,978 | | 429,978 | |
| Distressed firms | 30,900 | | 30,900 | |
| Pseudo R-squared | 0.153 | | 0.177 | |
| Log likelihood | -131,451.70 | | -127,673.50 | |
| Wald test | 49,903.0*** | | 55,783.8*** | |
| Likelihood ratio test | | | 7,556.43*** | |

Panel B. In-sample prediction tests

Hosmer-Lemeshow test: Percentage of the 30,900 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 9.70% | 9.17% |
| 8 | 12.06% | 11.86% |
| 9 | 19.54% | 18.67% |
| 10 | 46.33% | 47.01% |
| 8 to 10 | 77.93% | 77.54% |
| Area under the ROC curve | 0.8336 | 0.8482 |

Panel C. Out-of-sample prediction tests

A hold-out sample of 48,034 Italian, Portuguese and Spanish SMEs (195,236 firm-year observations) is used.

Hosmer-Lemeshow test: Percentage of the 3,434 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 9.55% | 9.11% |
| 8 | 11.68% | 10.80% |
| 9 | 19.60% | 19.57% |
| 10 | 46.48% | 47.38% |
| 8 to 10 | 77.75% | 77.75% |
| Area under the ROC curve | 0.8367 | 0.8497 |

## Table 8

## Group 3 (Czech Republic, Poland)

Panel A. Estimation results

The models are estimated for 2000-2009 data with yearly observations using the multi-period logit technique. All firm-specific variables are lagged by one year. The data set is limited to non-financial Czech and Polish SMEs. Parameter estimates are given first followed by chi-square values in parentheses. There are 48,470 firms in the sample (178,618 firm-year observations) out of which 4,278 distressed.

| | Group 3 Model I | | Group 3 Model II | |
|---|---|---|---|---|
| Earnings before taxes to total assets | -0.537*** | (-4.68) | -0.547*** | (-4.68) |
| EBITDA to interest expenses | -0.0000483*** | (-4.46) | -0.0000509*** | (-4.55) |
| Current liabilities to total assets | 1.312*** | (31.47) | 1.397*** | (32.84) |
| Cash flow to current liabilities | -0.315** | (-2.59) | -0.314** | (-2.64) |
| Turnover to total liabilities | -0.183*** | (-13.57) | -0.175*** | (-13.28) |
| Size (ln(totals assets)) | -0.120*** | (-10.16) | -0.0754*** | (-6.04) |
| Dummy equal to 1 if shareholders are more than 2 | -0.314*** | (-6.97) | -0.347*** | (-7.53) |
| Dummy equal to 1 if SME is located in an urban area | 0.319*** | (9.08) | 0.358*** | (9.92) |
| Duration | 0.363*** | (45.04) | 0.244*** | (25.89) |
| FX rate volatility | | | 122.6*** | (12.11) |
| 10-year government bond yield | | | 25.43*** | (14.63) |
| GDP growth (% change) | | | -11.62*** | (-22.52) |
| Industry 1 (Agriculture, Mining, Manufacturing) | | | | |
| Industry 2 (Transportation, Communication) | | | | |
| Industry 4 (Trade) | | | | |
| Industry 5 (Accommodation and Food) | | | | |
| Constant | -4.487*** | (-43.64) | -6.275*** | (-45.81) |
| Firm-year observations | 178,618 | | 178,618 | |
| Firms | 48,470 | | 48,470 | |
| Distressed firms | 4,278 | | 4,278 | |
| Pseudo R-squared | 0.214 | | 0.250 | |
| Log likelihood | -15,878.4 | | -15,147.9 | |
| Wald test | 8,206.4*** | | 8,083.9*** | |
| Likelihood ratio test | | | 1,460.92*** | |

* p<0.05, ** p<0.01, *** p<0.001

Panel B. In-sample prediction tests

Hosmer-Lemeshow test: Percentage of the 4,278 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 7.62% | 7.22% |
| 8 | 10.14% | 9.70% |
| 9 | 18.35% | 17.16% |
| 10 | 52.06% | 55.68% |
| 8 to 10 | 80.55% | 82.54% |
| Area under the ROC curve | 0.8653 | 0.8749 |

Panel C. Out-of-sample prediction tests

A hold-out sample of 5,340 Czech and Polish SMEs (19,844 firm-year observations) is used.

Hosmer-Lemeshow test: Percentage of the 427 distressed firms predicted in each decile in the year of their distress.

| Decile | | |
|---|---|---|
| 1 to 5 | 8.67% | 7.73% |
| 8 | 11.24% | 8.90% |
| 9 | 15.46% | 16.63% |
| 10 | 55.97% | 58.08% |
| 8 to 10 | 82.67% | 83.61% |
| Area under the ROC curve | 0.8632 | 0.868 |

## Table 9

**Subperiods' Analysis (8 countries)**

Panel A. Estimation results

The models are estimated over different subperiods (five-year rolling windows for 2002-2009 data) with yearly observations using the multi-period logit technique. Estimation results for the overall sample are also provided in the last two columns for comparison purposes (2000-2009 data). All firm-specific variables are lagged by one year. The data set includes SMEs from eight European economies. Parameter estimates are given first followed by chi-square values in parentheses.

| | 2002-2006 | | 2003-2007 | | 2004-2008 | | 2005-2009 | | Model III. | |
|---|---|---|---|---|---|---|---|---|---|---|
| Earnings before taxes to total assets | -0.819*** | (-9.21) | -0.764*** | (-9.02) | -0.824*** | (-11.07) | -0.757*** | (-14.16) | -0.763*** | (-15.53) |
| EBITDA to interest expenses | -0.0000248*** | (-5.91) | -0.0000390*** | (-9.78) | -0.0000477*** | (-12.25) | -0.0000544*** | (-15.24) | -0.0000451*** | (-14.58) |
| Current liabilities to total assets | 1.789*** | (68.72) | 1.684*** | (74.63) | 1.530*** | (75.25) | 1.379*** | (92.63) | 1.417*** | (102.97) |
| Cash flow to current liabilities | -0.557*** | (-5.66) | -0.635*** | (-6.49) | -0.493*** | (-5.87) | -0.523*** | (-9.10) | -0.491*** | (-9.16) |
| Turnover to total liabilities | -0.0900*** | (-13.08) | -0.0983*** | (-14.83) | -0.118*** | (-18.17) | -0.169*** | (-31.22) | -0.176*** | (-35.56) |
| Size (ln(totals assets)) | -0.0980*** | (-13.25) | -0.0316*** | (-4.85) | 0.0446*** | (7.34) | -0.0188*** | (-4.01) | -0.0913*** | (-22.14) |
| Dummy equal to 1 if shareholders are more than 2 | -0.245*** | (-11.31) | -0.279*** | (-15.03) | -0.246*** | (-14.68) | -0.277*** | (-20.73) | -0.272*** | (-21.76) |
| Dummy equal to 1 if SME is located in an urban area | 0.125*** | (4.93) | 0.0757*** | (3.53) | 0.0959*** | (5.24) | 0.141*** | (10.26) | 0.144*** | (11.01) |
| Duration | 0.358*** | (59.17) | 0.334*** | (86.41) | 0.238*** | (78.84) | 0.172*** | (75.43) | 0.228*** | (108.86) |
| FX rate (% change) | -1,421.8*** | (-29.32) | -1,452.5*** | (-33.02) | -478.9*** | (-13.02) | -541.8*** | (-14.16) | -1,689.9*** | (-59.01) |
| Unemployment | 2.117*** | (4.38) | -0.462 | (-0.97) | 2.082*** | (6.89) | 4.423*** | (28.04) | 1.914*** | (12.58) |
| Economic sentiment indicator | -0.0169*** | (-7.70) | -0.0368*** | (-20.08) | -0.0106*** | (-10.39) | -0.00570*** | (-6.75) | -0.0258*** | (-34.90) |
| Loans granted to non-financial sector (% change) | -6.238*** | (-50.60) | -5.288*** | (-48.89) | -2.347*** | (-20.75) | -4.202*** | (-49.85) | -4.407*** | (-58.07) |
| Years to resolve insolvency proceedings | -0.0497*** | (-5.03) | 0.0520*** | (8.94) | 0.0981*** | (23.87) | 0.157*** | (46.05) | 0.0958*** | (27.57) |
| Industry 1 (Agriculture, Mining, Manufacturing) | 0.0628* | (2.06) | 0.0211 | (0.80) | 0.0915*** | (3.75) | 0.0712*** | (3.82) | | |
| Industry 2 (Utilities, Transportation, Communication) | -0.186*** | (-4.13) | -0.0976** | (-2.65) | 0.0245 | (0.75) | 0.0112 | (0.46) | -0.0762*** | (-3.56) |
| Industry 3 (Construction) | -0.193*** | (-6.07) | -0.0267 | (-0.99) | 0.179*** | (7.32) | 0.218*** | (11.78) | 0.0798*** | (5.84) |
| Industry 4 (Trade) | -0.0571* | (-1.99) | -0.113*** | (-4.56) | -0.0202 | (-0.88) | 0.0265 | (1.56) | -0.0295* | (-2.50) |
| Industry 5 (Accommodation and Food) | 0.264*** | (5.57) | 0.156*** | (4.06) | 0.190*** | (5.82) | 0.226*** | (9.57) | 0.212*** | (10.18) |
| Constant | -3.896*** | (-16.69) | -1.952*** | (-9.26) | -5.317*** | (-42.28) | -4.978*** | (-53.39) | -2.523*** | (-30.26) |
| Firm-year observations | 1,079,429 | | 1,367,406 | | 1,704,810 | | 2,056,890 | | 2,721,861 | |
| Firms | 385,546 | | 637,299 | | 646,812 | | 636,008 | | 644,234 | |
| Distressed firms | 15,914 | | 20,665 | | 24,276 | | 42,351 | | 49,355 | |
| Pseudo R-squared | 0.200 | | 0.158 | | 0.150 | | 0.125 | | 0.171 | |
| Log likelihood | -58,989.5 | | -69,826.7 | | -91,025.0 | | -111,420.9 | | -204,538.30 | |
| Wald test | 23,784.5*** | | 31,818.7*** | | 39,646.7*** | | 68,451.6*** | | 84,526.8*** | |

* p<0.05, ** p<0.01, *** p<0.001

| | 2002-2006 | 2003-2007 | 2004-2008 | 2005-2009 | 2000-2009 |
|---|---|---|---|---|---|
| Panel B. Performance over next year | | | | | |
| Hosmer-Lemeshow test: Percentage of the distressed firms predicted in each decile in the year of their distress. | | | | | |
| Decile | | | | | |
| 1 to 5 | 8.05% | 11.84% | 11.94% | - | - |
| 6 | 4.96% | 5.25% | 6.12% | - | - |
| 7 | 8.83% | 8.71% | 9.01% | - | - |
| 8 | 14.43% | 17.56% | 13.18% | - | - |
| 9 | 19.67% | 20.34% | 18.99% | - | - |
| 10 | 44.06% | 36.30% | 40.75% | - | - |
| 8 to 10 | 78.15% | 74.20% | 72.93% | - | - |
| 9 to 10 | 0.8177 | 0.7825 | 0.7963 | - | - |
| Area under the ROC curve | | | | - | - |
| Panel C. Performance over last year (2009) | | | | | |
| Hosmer-Lemeshow test: Percentage of the distressed firms predicted in each decile in the year of their distress. | | | | | |
| Decile | | | | | |
| 1 to 5 | 12.80% | 12.25% | 11.94% | - | - |
| 6 | 6.27% | 6.20% | 6.12% | - | - |
| 7 | 10.28% | 9.63% | 9.01% | - | - |
| 8 | 16.34% | 16.19% | 13.18% | - | - |
| 9 | 18.45% | 18.91% | 18.99% | - | - |
| 10 | 35.86% | 36.82% | 40.75% | - | - |
| 8 to 10 | 70.65% | 71.93% | 72.93% | - | - |
| 9 to 10 | 0.7795 | 0.7852 | 0.7963 | - | - |
| Area under the ROC curve | | | | - | - |

## Appendix A. List of Systematic Variables

The appendix provides a list of the systematic variables that we examine, and their expected signs, calculation methods, lags and data sources.

| | |
|---|---|
| **Business cycle** | |
| Change of the exchange rate | (-) Raw data are daily. We calculate the average daily change of the USD/EURO (for Eurozone members) and of USD/national currency (for non-Eurozone members) for the year before the closing. We do not lag this variable as data are accessible on real time. Source: European Central Bank. |
| Debt as a percentage of the GDP | (+) Raw data are quarterly. We take the average percentage over a four quarter period before the closing. We lag this variable by two quarters. Source: Eurostat. |
| Disposable income growth | (-) Raw data are quarterly. We take the disposable income change between the four quarters before the closing and the corresponding four quarters of the previous year. We lag this variable by one quarter. In the Eurostat data, year 2005 is used as the reference to measure disposable income at constant prices. Figures are also seasonally adjusted and adjusted by working days. Source: Eurostat. |
| Economic sentiment | (-) Raw data are monthly. This indicator is calculated by the Directorate General of Financial Affairs of the European Commission. It is calculated as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program. We take the average of the twelve months before the closing. We lag this variable by one month. Source: Eurostat. |
| GDP growth | (-) Raw data are quarterly. We take the GDP percentage change between the four quarters before the closing and the corresponding four quarters of the previous year. We lag this variable by one quarter. In the Eurostat data, year 2005 is used as the reference to measure GDP at constant prices. Figures are also seasonally adjusted and adjusted by working days. Source: Eurostat. |
| Inflation | (+) Raw data are monthly. We take the annual rate of change of the Harmonized Index of Consumer Prices (HICP). Specifically, we calculate the change of the index between the closing month and the corresponding month of the previous year. We lag this variable by one month. Source: Eurostat |
| Oil price | (+) Raw data are monthly (historical close). We take average of the one month forward prices of brent crude oil for the twelve months before the closing. We do not lag this variable as data are accessible on real time. Source: European Central Bank. |
| Surplus/deficit as a percentage of the GDP | (-) Raw data are quarterly. We take the average percentage over a four quarter period before the closing. We lag this variable by two quarters. Source: Eurostat. |

| | |
|---|---|
| Unemployment | (+) Raw data are monthly. We take the average harmonized unemployment rate (International Labor Organization definition) over a twelve month period before the closing. We lag this variable by one month. Source: Eurostat. |
| Volatility of the exchange rate | (+) Raw data are daily. We calculate the volatility of the daily change of the USD/EUR (for Eurozone members) and of USD/national currency (for non-Eurozone members) for the year before the closing. We do not lag this variable as data are accessible on real time. Source: European Central Bank. |
| **Credit conditions** | |
| 10-year government bond yield change | (+) Raw data are monthly. We take the annualized 10-year government bond yield (Maastricht definition) of the closing month. We do not lag this variable as data are accessible on real time. Source: Eurostat. |
| Bank lending to the non-financial sector | (-) Raw data are monthly. We take the percentage change between the closing month and the corresponding month of the previous year. We lag this variable by one month. Source: Datastream. |
| **Financial market** | |
| Stock index return | (-) Raw data are monthly. We take the one year return of the national stock market index (change between the closing month and the corresponding month of the previous year). We do not lag this variable as data are accessible on real time. Source: Eurostat. |
| **Insolvency codes** | |
| Recovery rate | (-) Raw data are annual. This indicator is calculated by the World Bank under the Doing Business project and measures the percentage that claimants (creditors, tax authorities, and employees) recover from an insolvent firm for each country. We lag this variable by one year. Source: World Bank. |
| Time to resolve insolvency proceedings | (+) Raw data are annual. This indicator is calculated by the World Bank under the Doing Business project and measures the number of years from the filing for insolvency in court until the resolution of distressed assets for each country. We lag this variable by one year. Source: World Bank. |

## Appendix B. Insolvency Regimes

The appendix provides an overview of the insolvency regimes in the countries of our study. The first column gives the average percentage that claimants recover form an insolvent firm in the years 2000-2009, the second column measures the average years from the insolvency filing until the resolution of assets in the same period and the third column is the ratio of the two previous columns. Data are from World Bank and authors' calculations.

| | Recovery rate (%) | Years to resolve insolvency | Recovery rate per year (%) |
|---|---|---|---|
| Italy | 48.22 | 1.80 | 26.79 |
| Portugal | 73.23 | 2.00 | 36.62 |
| Spain | 72.90 | 1.50 | 48.60 |
| France | 46.19 | 1.90 | 24.31 |
| Germany | 82.32 | 1.20 | 68.60 |
| United Kingdom | 85.31 | 1.00 | 85.31 |
| Czech Republic | 17.23 | 8.39 | 2.05 |
| Poland | 32.31 | 3.00 | 10.77 |