Journal of Banking & Finance 64 (2016) 112–135

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Forecasting distress in European SME portfolios ☆

Sara Ferreira Filipe, Theoharry Grammatikos, Dimitra Michala *

Luxembourg School of Finance, University of Luxembourg, 4 rue Albert Borschette, L-1246 Luxembourg, Luxembourg

ARTICLE INFO

Article history:
Received 8 April 2014
Accepted 23 December 2015
Available online 2 January 2016

JEL classification:
C13
C41
C53
G33

Keywords:
Credit risk
Distress
Forecasting
SMEs
Logit

ABSTRACT

In this paper, we examine idiosyncratic and systematic distress predictors for small and medium sized enterprises (SMEs) in Europe over the period 2000–2009. We find that SMEs across European regions are vulnerable to common idiosyncratic factors but systematic factors vary. Moreover, systematic factors move average distress rates and small SMEs are more vulnerable to these factors compared to large SMEs. By including many very small companies in the sample, our models offer unique insights into the European small business sector. By exploring distress in a multi-country setting, the models uncover regional vulnerabilities. Finally, by incorporating systematic dependencies, the models capture distress co-movements.

1. Introduction

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 (EU) in particular, SMEs represent 99% of all enterprises and contribute to more than half of all value-added created by businesses. Despite their importance, SME credit risk remains largely unexplored by the academic literature, mainly due to the lack of appropriate data.

In this paper, we explore a dataset that is representative of the European SME sector because it includes a high number of very small companies. This is important for Europe, where nine out of ten SMEs have fewer than 10 employees and turnover up to €2million. To our knowledge, we are the first to examine distress in a multi-country setting, since earlier studies always focus on a single economy. Hence, we are able to uncover regional vulnerabilities, perform comparisons and study the need of regional models in international SME portfolios. In addition to idiosyncratic distress determinants, we consider systematic factors, such as the macroeconomy, bank lending conditions, and legal aspects. Therefore, we are able not only to compute individual distress probabilities, but also to estimate average distress rates in the economy and capture distress co-movements.

Our paper 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 face more challenges in their credit risk modeling. In fact, widely used structural market-based models, such as the distance-to-default (*DD*) measure inspired by Merton (1974), cannot be applied to the non-listed SME setting due to the unavailability of market data. Instead, empirical predictive models such as credit scoring approaches (i.e. Altman, 1968; Edminster, 1972) are most commonly use. Many authors, such as Dietsch and Petey (2004), Berger and Udell (2006), and Beck et al. (2008), note the need for SME specific research. In line with their concerns, Altman and Sabato (2007) (in an early SME study) develop a

☆ We would like to thank Edward Altman, Jos Van Bommel, Michel Dietsch, Federico Galizia, Jean-Daniel Guigou, Helmut Kraemer-Eis, Edward Olding, the Editor and two anonymous referees as well as participants in the 6th Annual Meeting of the EIB Institute's Knowledge Program, the LSF Internal PhD Seminar Series, the XIV 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.

* Corresponding author. Tel.: +352 661332225.

E-mail addresses: sara.ferreira@uni.lu (S.F. Filipe), theoharry.grammatikos@uni.lu (T. Grammatikos), dimitra.michala@gmail.com (D. Michala).

http://dx.doi.org/10.1016/j.jbankfin.2015.12.007
0378-4266/

one-year default prediction model using only accounting information. They apply panel logit estimation on a sample of around 2000 US SMEs over the period 1994–2002. They find that their model outperforms generic corporate models such as Altman's Z"-score (Altman and Hotchkiss, 2005).

Stein (2002), Grunert et al. (2005) and other authors note the possibility of using qualitative variables in default prediction models to improve discrimination. In the specific case of SMEs (where there is usually a problem of scarcity of reliable "hard" financial information) such non-financial elements can be very useful when trying to predict distress. Altman et al. (2010) combine both qualitative and financial information in a default prediction model for SMEs in the UK. They find that data relating to legal action by creditors, company filings and audit reports/opinions significantly increase the performance of their model. However, such information is not always available sufficiently in advance to facilitate timely predictions.

Another strand of literature (though not focusing on SMEs) analyzes the additional benefit of using macroeconomic variables to forecast distress. Two influential US studies of this nature are Duffie et al. (2007) and Campbell et al. (2008). International studies always focus on specific countries, such as Jacobson et al. (2005), Carling et al. (2007) and Jacobson et al. (2013) on Sweden, Bonfim (2009) on Portugal, Bruneau et al. (2012) on France, and Nam et al. (2008) on Korea. The above articles find that macroeconomic variables are important for explaining the time-varying default likelihoods, but they examine relatively larger (and, in the case of US, listed) corporates. The authors also note the importance of industry effects. For instance, Chava and Jarrow (2004) observe improving forecasting performance by including industry groupings in their models.

Our paper is related to the studies of Glennon and Nigro (2005), Altman et al. (2010), Jacobson et al. (2013) and Laerkholm-Jensen et al. (2015), who, respectively, examine business cycle effects on SMEs defaults in the US, the UK, Sweden, and Denmark. Glennon and Nigro (2005), using a dataset of US small loans, include business cycle dummy variables, industrial production index growth and rates of regional business bankruptcies. They find that the failure of a small loan is closely related to both regional and industrial economic conditions. Altman et al. (2010) use sector-level failure rates of SMEs in the UK and also report a significant relationship with failure probability. Jacobson et al. (2013) consider both idiosyncratic and macroeconomic factors for the entire Swedish corporate sector and perform a careful cross-industry comparison. Finally, in a recent working paper, Laerkholm-Jensen et al. (2015) find that macro variables play the most important role in default prediction over time for their Danish sample. Our paper extends the above studies by using a wider sample that includes SMEs from different European countries, by allowing for regional models and comparisons, and also by examining a larger variety of systematic factors (ranging from exchange rates to bank lending conditions). Europe offers a unique setting for such a study compared to the US due to the higher level of variation between economic and legal environments faced by SMEs. Another slight complication of the US studies is that they often use the average sample default rate as an explanatory variable in their models in order to capture business cycle effects. This technique can introduce bias and may result in opposite coefficient signs (Gormley and Matsa, 2014).

In our study, we find that, in addition to indicators of profitability, coverage, leverage and cash flow, the location and the number of shareholders are important distress determinants for SMEs. We also confirm that systematic factors significantly affect average distress rates in the European economy, a finding that is well-documented in previous US and international literature (Duffie et al., 2007; Carling et al., 2007; Altman and Rijken, 2011; Jacobson et al., 2013; Laerkholm-Jensen et al., 2015). Nevertheless, industry effects often do not demonstrate significance. Moreover, we examine interaction effects between SMEs' size and systematic variables. We find that as SMEs become larger, they are less vulnerable to systematic factors, a finding that is particularly important in light of the current Basel regulations.

Our most interesting results appear when we split our sample into regional sub-samples. We find that SMEs in different regions are vulnerable to the same idiosyncratic factors but coefficient levels differ among regions. Most importantly, SMEs in different regions are exposed to different systematic factors, according to region-specific conditions and characteristics. Our regional distress models always perform better than a generic unrestricted model estimated for each regional sub-sample and a generic restricted model that imposes the same parameters to the systematic factors in all regional sub-samples. These findings indicate the importance of using regional models for distress prediction in international SME portfolios. Finally, our results remain robust to different distress definitions, estimation techniques and time periods.

The paper is organized as follows: Section 2 describes the methodology and the reasons for its selection. Section 3 describes the dataset, discusses the choice of variables and presents summary statistics. Section 4 presents the models and discusses the results, Section 5 presents the robustness tests, and Section 6 concludes.

2. The methodology

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

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

where $Y_{i,t}$ is an indicator that equals one if the firm i 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 γy_{t-1} is the baseline hazard function that includes some other time-dependent variables y_{t-1}.

As outlined above, in this paper we apply a reduced method and use a distress indicator dummy variable ($Y_{i,t}$) as the dependent variable in our logit regression. Another option would be to estimate probabilities of distress using a structural method, such as the Merton approach or credit spreads, and then use these estimated probabilities as the dependent variable in our regression. Let us elaborate why this second option is not applicable in the case of non-listed SMEs. First, the Merton method requires daily data on stock prices in order to find the asset volatility of the company. Our sample though does not have traded companies but only private ones, since the vast majority of European SMEs are small and do not satisfy entry requirements to stock exchanges, thus there are no market data available for them. Similarly, Allen et al. (2004) study credit risk modeling techniques for small business loans and write that: "*The problem with such models (Merton based models such as KMV) is that retail borrowers often do not have publicly traded stock, and therefore, equity prices may not be available or may be unreliable because of liquidity problems*". They also write that "*Merton (1974) [...] uses equity volatility to estimate asset volatility, since both the market value of firm assets and asset volatility are unobservable. Retail clients do not have a series of equity prices that can be used to estimate asset values or asset volatility*". Second, credit spreads (CDS or bonds) cannot be used here, since they require other market data which are also not available for such small companies. The companies in our sample do not have CDS information and do not issue bonds, as their main funding sources are own resources and bank debt.

Back to Eq. (1), the baseline hazard y_{t-1} 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.

Shumway (2001) proves that, for a discrete random variable t, a multi-period logit model is equivalent to a discrete-time hazard model with an adjusted standard error structure. We need to adjust the standard errors because test statistics produced by the logit program assume that the number of independent observations is the number of firm-years and they also ignore the panel structure of the data. Calculating correct test statistics requires the adjustment of the sample size to account for dependence among firm-year observations. The firm-year observations of a particular firm cannot be independent, since a firm cannot fail in period t if it failed in period $t-1$. Likewise, a firm that survives to period t cannot have failed in period $t-1$. Thus, the correct value of n for test statistics is the number of firms in the data, not the number of firm-years. The χ^2 test statistics produced by the logit program are of the form:

$$\frac{1}{n}(\hat{\mu}_k - \mu_0)'\Sigma^{-1}(\hat{\mu}_k - \mu_0) \sim \chi^2(k), \quad (2)$$

where there are k estimated moments being tested against k null hypotheses, μ_0. Dividing these test statistics by the average number of firm-years per firm makes the logit program's statistics correct. This is equivalent to calculating firm clustered-corrected standard errors to adjust for the number of firms in our samples. Specifically we use Huber/White standard errors (calculated from Huber/White sandwich covariance matrix, see Froot, 1989; White, 1994; Wooldridge, 2002).

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 and consequently, 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 follow a technique similar to Carling et al. (2007) and introduce one more factor, the "duration" variable that accounts for the "time-at-risk" of firms only during the sample period, (i.e. the number of years that a firm stays in the sample). The value of this variable is given by the formula duration = t 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, 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. This is because, 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.[1]

3. 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 (purchased from Bureau Van Dijk) to detect the status of each firm in each year and to 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.

3.1. Definition of distress

We classify firm-years into two mutually exclusive categories: "distressed" and "healthy". A firm-year is *distressed* if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) appears with one of the following statuses "defaulted", "in receivership", "bankrupt", "in liquidation" or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. A firm-year is *healthy* in all other cases. Specifically, we consider as healthy: (i) firm-years of distressed companies before the last available firm-year; (ii) all firm-years of firms that disappear from the sample for a specified reason other than distress (i.e. merger or acquisition); (ii) all firm-years of firms that have no updated status information and disappear from the sample before 2010 without negative equity in the last year; (iv) all firm-years of firms that remain active until 2010.

Let us elaborate further on the above. Firms enter the sample anytime during the years 1999–2008. We track them until 2010 and use financial statements from the years 1999–2008 to predict distress on a one-year horizon for the period 2000–2009.[2] There are two cases:

Case 1: Firms that remain active in the sample until 2010 (in the sense that they report financial statements until 2010). All firm-year observations for these firms for the estimation period 2000–2009 are classified as healthy.

Case 2: Firms that disappear from the sample (in the sense that they no longer report financial statements) earlier than 2010. For these firms either we consult the available status information to find out why they disappeared or, when no updated status information is available, we consider as distressed the last available firm-year when the disappearing firm has negative equity in this particular year.

It is important to note that the negative equity condition is not used for any of the firm-years of case 1. Now we explain the reason why we add this condition. Our intention in this study is to proxy for distress and not only failure. Thus, we are not only interested in incidents that are strictly determined by legal insolvency procedures. The extended indicator that we use is more appropriate for SMEs because these companies often do not follow such legal procedures at all. A characteristic example is Italy, where there is no clear 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 (e.g. Gilson and Vetsuypens (1993) show in a US study that many filings are missing for bankrupt firms). When these procedures are mandatory, legal insolvency is often related to negative equity. For example, in Germany, firms are

[1] However, the "duration" variable is still an imperfect measure. This is because we can underestimate the lifespan of firms that default in the beginning of the sample period.

[2] Although we have financial statements data for 2009 and 2010, we do not use them to predict distress for 2010 and 2011 because we do not know which SMEs become distressed during these years. The negative equity condition does not help in this case, since the last available year of our sample is 2010 and we do not know which SMEs disappear the following year, thus we cannot construct our distress indicator for 2010.

obliged to file for bankruptcy once their equity turns negative (Davydenko and Franks, 2008) and in France even earlier, when their equity drops below a certain threshold (LaPorta et al., 1998).

As a result, the proper tracking of the status of SMEs and their distress rates is a very challenging task. There are many different reasons for which an SME can go out of business but owners rarely report these reasons and authorities rarely document them. Watson and Everett (1993) find that small businesses often close for reasons other than distress. For example, a small business can be successful but the owner may still close it voluntarily to accept employment with another company or retire. Headd (2003) finds that only one out of three of start-ups close under conditions that the owners consider unsuccessful. The Amadeus and Orbis databases cooperate in different countries with credit bureaus which provide firm status information. In around 40% of cases though, a firm disappears from the database but the status information remains outdated. In order to separate the cases of closure from the ones of distress for these firms, we need to make a reasonable assumption. This is why we add the negative equity condition.

This condition is well-rooted in various academic studies. A large strand of literature links equity values with firm distress. For example, Davydenko (2012) describes as economic default the point when a firm's equity turns negative and characterizes this as a distress-triggering event. The definition in Chapter 7 of the US Bankruptcy Law is very similar. Davydenko (2012) finds support for models in which the default timing is chosen endogenously such as Merton's *DD*. Ross et al. (2010) point out that a stock-based insolvency occurs when a company has negative equity.

In our sample, we observe that negative equity is 200% more likely for firms that disappear from the sample at some point before 2010 than for firms that remain active. From an accounting perspective, negative equity is almost always connected with accumulation of past losses. From a capital structure point of view, negative equity means that the company's total liabilities are higher than its total assets. In both cases, a negative value for equity is a flag that the company is undergoing serious financial difficulties and it is a good proxy for distress (and not only failure).

To verify our point, in Section 5.1 we perform two robustness tests, in each one applying an alternative distress definition. In the first alternative distress definition, we exclude all firms that disappear from the sample before 2010 without updated status situation. These include firms that, under the main distress definition are classified as distressed if their equity is negative in the last year. Thus, the first alternative distress definition is strictly linked to a legal insolvency procedure. In the second alternative distress definition, we exclude all firms that have negative equity in one or more of the years of their existence in the sample. Thus, under the second alternative distress definition, we essentially lose some distress-related information because we include in the sample only firms with non-negative equity. Despite the fact that our sample size decreases significantly in both cases, our results remain robust. In Section 5.1 and Appendix C, we report the estimation results and comparative statistics for our two alternative distress definitions versus our main distress definition. Finally, in unreported results, we replace the negative equity condition with one for negative earnings before interest, taxes, depreciation and amortization (EBITDA). EBITDA is often used in the academic literature as a proxy for operating cashflow. In our sample, 67% of SMEs that have negative equity also report negative EBITDA in the same year. Our results also remain substantially similar under this distress definition as well.

3.2. Sample selection

In our sample SMEs come from eight European countries, namely Czech Republic, France, Germany, Italy, Poland, Portugal, Spain and the United Kingdom. We select these countries for two reasons: (i) our version of Amadeus and Orbis Europe databases has only European coverage but data are not of the same quality for all countries. (For Scandinavian countries there are very few distress cases for estimation purposes and for most eastern European countries there are very few firms with complete information); (ii) this particular set of countries creates a combination that reflects the variability of SMEs across the EU. This is obvious from Table 1, which provides an overview of the key indicators for SMEs in the EU27 and in the countries of our sample. In Italy, Portugal and Spain, SMEs account for larger than EU-average shares of total employment and value added, and present in higher density. This suggests that SMEs in these economies have a greater role than in most EU countries. On the other hand, for France, Germany and the UK, these figures are consistently lower than the EU average. For the Czech Republic and Poland, the share of employment and value added for SMEs is similar.

Table 1
Key indicators. The economic and social contribution 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 1000 inhabitants. Data are from Eurostat.

	(%) 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

To study whether the distress determinants differ across Europe and in order to perform comparisons, we split our sample in regional sub-samples. We select the groups based on the following criteria: (ii) the importance of SMEs in the local economies, reflected in Table 1; (ii) geography, i.e. west, south, east; (iii) the similarity of the macroeconomic environment, i.e. correlations of macroeconomic variables, currency etc; and lastly, (iv) previous literature. Thus, we form three groups. Group 1 includes the relatively stronger economies of western Europe, namely France, Germany and the UK, group 2 includes the peripheral economies of southern Europe, namely Italy, Portugal and Spain, and group 3 includes two economies from eastern Europe, namely the Czech Republic and Poland. We discuss criterion (i) above and criterion (ii) is clear. Concerning criterion (iii), when we calculate correlations of macroeconomic variables between all country combinations, we find a clear division along the regions. Finally, on criterion (iv), these countries are often bundled together in existing studies (Jaumotte and Sodsriwiboon, 2010; Grammatikos and Vermeulen, 2012; Perego and Vermeulen, 2013).

Because of the European focus of the study, we adopt the European Commission's definition for SMEs (European Commission, 2007), 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 fewer than 250 employees and, either, annual turnover up to €50 million, or total assets up to €43 million; (ii) no single company holds more than 25% of their equity; (iii) they do not have subsidiaries; (iv) they have up to ten shareholders; (v) they have at least two years of data available; (vi) they are not firms in the financial sector.

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 sub-samples. It gives the total number of SMEs at the beginning of the year, the number of distressed SMEs during the year and the distress rate per year. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain, and group 3 includes Czech Republic and Poland.

Year	Overall sample			Group 1			Group 2			Group 3		
	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	2221	1	0.05
2002	204,531	3802	1.86	99,815	2125	2.13	100,466	1649	1.64	4250	28	0.66
2003	194,768	5961	3.06	91,761	4003	4.36	94,857	1935	2.04	8150	23	0.28
2004	146,877	1250	0.85	52,031	865	1.66	81,727	331	0.41	13,119	54	0.41
2005	167,837	1403	0.84	53,609	822	1.53	99,053	377	0.38	15,175	204	1.34
2006	256,732	1873	0.73	70,242	902	1.28	164,105	734	0.45	22,385	237	1.06
2007	463,732	8134	1.75	95,393	1600	1.68	331,731	5932	1.79	36,608	602	1.64
2008	498,358	9194	1.84	88,606	1427	1.61	369,487	6977	1.89	40,265	790	1.96
2009	463,652	17,546	3.78	75,065	2248	2.99	352,923	12,959	3.67	35,664	2339	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	4278	2.40

We need criteria (ii)–(iv) to ensure that the companies are independent.[3] Specifically, since we cannot track the subsidiaries and check if the companies still satisfy the criteria to be classified as SMEs once they become subsidiaries, 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 they are possibly outliers. As to criterion (v), 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. Finally, on criterion (vi), we follow Shumway (2001) and other authors and exclude financial firms from the sample (NACE[4] rev.2 codes from 64 to 68) due to financial firms having reporting practices that preclude combining them with other firms in models using financial information.

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). These constrains limit our initial dataset by around 25%.

Our estimation sample consists of 2,721,861 firm-years observations (644,234 firms) out of which 49,355 are distressed. We additionally 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 5487 are distressed. 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 relatively high 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. 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. 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).

3.3. 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.

3.3.1. 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 employee efficiency.[5] 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 ninety-ninth percentile, a popular technique known as winsorizing. Finally, because annual reports for SMEs become available with a significant time delay, we lag all ratios in the estimations by 12 months. 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 apply the area under the curve (AUC) criterion. The AUC is constructed from the estimated distress probabilities versus the actual status of the firms

[3] 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 during periods of poor financial performance. Moreover, the group may have reasons to support a subsidiary other than for financial reasons. Finally a subsidiary may be in distress as a result of group-wide distress.

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

[5] 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.

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 healthy years, and Panel C describes the distressed years. Panels D, E and F describe the distributions for Groups 1, 2 and 3 respectively. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain, and group 3 includes Czech Republic and Poland. The sample period is from 2000 to 2009. All ratios are winsorized 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. Overall sample					
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	2927.14	0.34	0.86	4.13
Min	−0.85	−2600.00	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 years					
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	2945.99	0.33	0.86	4.15
N: 2,672,506					
Panel C. Distressed years					
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	1448.37	0.56	0.59	2.50
N: 49,355					
Panel D. Group 1					
Mean	0.08	1064.80	0.61	0.32	3.76
Median	0.06	12.75	0.60	0.16	3.18
Std.Dev.	0.17	3682.35	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	2426.19	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	3357.95	0.41	1.19	6.39
N: 178,618					

in each year for all possible cut-off probability values. Thus, we 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 multi-collinearity 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 significant in previous studies. Finally, we apply a forward stepwise selection procedure of the remaining ratios, setting the significance level at 10% and performing the likelihood ratio test which is more accurate than the standard Wald test.

Table 3 reports summary statistics for the five ratios that are found to be the most effective in predicting distress. A comparison of Panels B and C in Table 3 reveals the differences between 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 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 depending on the region.

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 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 the 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

Table 4

SMEs by size and industry. (A) The classification of SMEs by size. The first column shows the size classes. The second column shows the firm data available in each class, the third column shows the percentage of firm data available in each class, the fourth column shows the number of firm-years data available in each class and the fifth column shows the distressed firm-years data available in each class. Finally the sixth column shows the distress rate as a percentage of total firm-years in each class. The sample period is from 2000 to 2009. (B) The classification of SMEs by wide industry sectors. The first column shows the sectors. The second column shows the firm data available in each sector, the third column shows the percentage of firm data available in each sector, the fourth column shows the number of firm-years data available in each sector and the fifth column shows the distressed firm-years data available in each sector. Finally the sixth column shows the distress rate as a percentage of total firm-years in each sector. The sample period is from 2000 to 2009.

Size					Firms	(%) firms	Firm-years	Distressed	(%) distressed
Cat.	Employees	Turnover	or	Assets					
Panel A. Size classification									
Medium	<250	⩽€50 m		⩽€43 m	21,408	3.32	123,123	1815	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)									
Sector					Firms	(%) firms	Firm-years	Distressed	(%) distressed
1. Agriculture, Mining and Manufacturing					133,746	20.76	608,696	9815	1.61
2. Transportation, Communication and Utilities					45,413	7.05	182,180	2827	1.55
3. Construction					113,147	17.56	482,031	9170	1.90
4. Trade					214,061	33.23	946,368	16,291	1.72
5. Accommodation and Food					36,235	5.62	128,225	3691	2.88
6. Other services					101,632	15.78	374,361	7561	2.02
Total					644,234	100.00	2,721,861	49,355	1.81

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 we get insignificant results for more detailed industry classifications. Panel B of Table 4 shows the way these broad sectors are partitioned. This initial evidence shows that Accommodation and Food has the highest distress rate and Transportation, Communication and Utilities the lowest.

Finally, we include a dummy for shareholders (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 injections of capital in difficult times, thus will 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 in monitoring firm performance compared to unlimited partners, leading limited SMEs to distress more frequently. Whereas, (as we show in the "results" section) 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.

Lastly we examine the age of a smaller sample of firms for which we have the date of establishment. 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 (but not monotonically) related to distress.

3.3.2. 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 calculate averages. We also usually lag them in order to avoid causality considerations and because they are available for forecasting with a time delay. 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).

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 the appreciation of the local currency, disposable income, 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. 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 a 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 uncertainty about future economic conditions and should be positively related to distress.

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 rates makes it harder for SMEs to borrow, whereas higher bank lending growth results in greater access to finance.

Finally, at this point, we need to elaborate on the effect of 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 default outcomes. 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. In general, 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 this process. This is also obvious in Appendix B. Countries where the insolvency procedure takes longer (such as the Czech Republic and Poland) score very low as regards the percentage of recoveries. The opposite is true for countries with swift procedures, 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, those which significantly influence the probability of distress for SMEs, we follow a standard procedure. First, we fit the models using only accounting information. 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 sub-samples, 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, lead to unreasonable signs of the estimated coefficients, and create 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 sub-samples, we anticipate that systematic variables will vary across regions. Based on economic intuition, we suspect that group 2 (Italy, Portugal, Spain – the peripheral economies of south Europe) is more exposed to the macroeconomic situation compared to group 1 (France, Germany, United Kingdom – more stable economies). Also, we suspect that group 3 (Czech Republic, Poland) is exposed to additional currency risk since these countries are not members of the Eurozone.

Finally, we also 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 macroeconomic situation.

4. The results

In this section, we present results of (i) models fitted and estimated using the overall sample, (ii) models fitted using the overall sample and estimated using the regional sub-samples, (iii) models that impose the same parameters to the systematic variables in all regional sub-samples but where the firm-specific coefficients are estimated for each regional sub-sample, and (iv) models fitted and estimated using the regional sub-samples. We refer to the first three classes of models as generic models, and the fourth class of models as regional models. We identify interesting differences among the European regions, and we show that regional distress models always perform better than a generic unrestricted model estimated for each regional sub-sample and a generic restricted model that imposes the same parameters to the systematic variables in all regional sub-samples.

4.1. Generic models estimated for the 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 a dummy for SMEs in urban areas), model II includes both the idiosyncratic and the systematic variables, model III also includes 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.

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 is positively related to leverage (current liabilities to total assets). Surprisingly, we do not find liquidity ratios as significant in the models. An explanation is that information contained in these ratios is proxied by others. That is, 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). In unreported results, we also test for the non-linear effects of size, by introducing the natural logarithm of squared total assets. We find a positive coefficient, indicating that for the largest SMEs distress risk starts to increase, probably because these companies are more likely to be pursued in liquidation process by their creditors. Two additional interesting findings in accordance with our predictions are that SMEs with less than three shareholders and SMEs in urban areas on average face higher risks. 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 risks faced by SMEs in urban areas is that these companies face higher competition (due to geographical proximity) and pay higher rent than their counterparts in non-urban areas. Another reason may be that owners of urban SMEs are less willing to support their enterprises in times of difficulties. This strategic distress caused because it is a more often a viable option for business owners to close the business and find employment elsewhere. [6] These effects seem to outweigh the fact that there is a larger customer base available for urban SMEs.

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. As we hypothesized, an appreciation of the currency, an increase in the economic sentiment indicator and an increase in lending by banks result in lower distress rates. Conversely, an increase in unemployment and a greater number of years to achieve insolvency

[6] Dietsch and Petey (2004) show something similar. Specifically, they find evidence from French SMEs that more attractive and wealthy regions demonstrate higher distress rates on average.

Table 5
Generic models estimated for the overall sample. The models are estimated for 2000–2009 with lagged yearly observations using the multi-period logit technique. The data-set includes non-financial SMEs from eight European economies, namely France, Germany, the UK, Italy, Portugal, Spain, Czech Republic and Poland. There are 644,234 firms in the sample (2,721,861 firm-year observations) out of which 49,355 distressed. Standard errors are firm clustered-corrected (Huber/White standard errors). Parameter estimates are given first followed by chi-square values in parentheses. * denotes significance at a 5% level, ** at a 1% level and *** at a 0.1% level. FX rate is calculated in relation to the US dollar (USD/national currency). Unemployment is measured in %. The economic sentiment indicator is calculated by the Directorate General of Financial Affairs of the European Commission as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program.

	Model I		Model II		Model III		Model IV		Model V	
Panel A. Estimation results										
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)
FX rate (% change)			−1686.8***	(−59.04)	−1689.9***	(−59.01)	−2627.20***	(−68.63)	−2695.82***	(−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)							1796.63***	(35.16)	1737.00***	(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)							1936.71***	(20.51)	1975.46***	(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)

Table 5 (*continued*)

	Model I	Model II	Model III	Model IV	Model V	
Age					−0.0133***	(−17.30)
Dummy equal to 1 if age is between 3 and 9 years					0.5501***	(43.76)
Constant	Yes	Yes	Yes	Yes	Yes	
Duration	Yes	Yes	Yes	Yes	Yes	
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***	
Panel B. In-sample prediction tests						
Hosmer–Lemeshow decile						
1–5	11.38%	11.09%	10.96%	10.67%	10.24%	
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%	
8–10	75.83%	76.59%	76.66%	77.06%	77.16%	
Area under the ROC curve	0.824	0.838	0.839	0.843	0.857	
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 decile						
1–5	11.46%	11.35%	11.26%	10.30%	10.15%	
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%	
8–10	75.54%	76.27%	76.23%	77.55%	76.82%	
Area under the ROC curve	0.823	0.837	0.837	0.844	0.847	

resolution result in higher distress rates. 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) 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 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 we 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). For example, 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 has less influence on their distress probability (but to the same direction) compared to micro firms. There are similar patterns with all other interaction effects with the exception of unemployment. Interestingly, the additional effect of unemployment for small and medium firms is of a higher magnitude (−10.495 and −11.241 respectively) in absolute terms than the coefficient for unemployment (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 the slightly smaller sample for which we have available data on age. We find, in accordance with previous literature, that older firms are safer. We also follow Altman et al. (2010) and check for non-linear effects of age. As in their study, we find a positive and statistically significant coefficient for a dummy variable equal to one if SMEs are between three and nine years old.

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). In order to evaluate more closely the performance of our models, we perform in-sample and out-of-sample testing. We

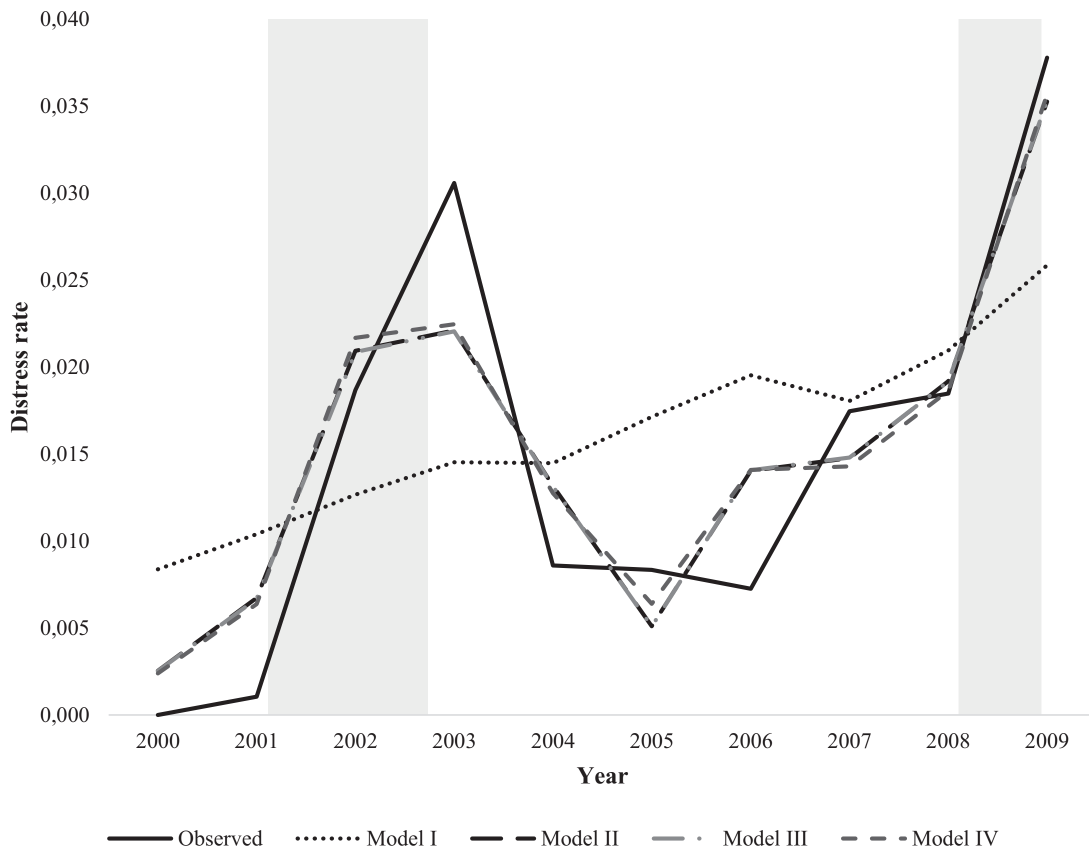

Fig. 1. Predicted and observed distress rates. The figure plots the predicted distress rate based on models I–IV of Table 5, along with the observed distress rate, for the period 2000–2009. The columns denote recession periods in the euro area, as indicated by OECD.

employ two widely used measures, the Hosmer and Lemeshow grouping based on estimated distress probabilities and the area under the curve (AUC).

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 the 2000–2009 period and calculate the corresponding percentages of the distressed firms in each decile.

The 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 0.7–0.8 represents good discriminatory power, an AUC of 0.8–0.9 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. Several statistics are equivalent to the AUC, such as the accuracy ratio, the Gini coefficient and the Mann–Whitney–Wilcoxon test (Engelmann et al., 2003). While the Hosmer and Lemeshow method assesses mainly calibration, the AUC assesses discrimination.

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 0.824 to 0.838. This result is stronger 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 (76.59–76.66%). 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 factors, we notice a further increase in the percentage of distressed firms in the last three deciles (76.66–77.06%). AUC also increases from 0.839 to 0.843. Moving to model V, it seems that age also helps slightly. However, we cannot 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. As seen, all results follow the same patterns as for the in-sample tests.

The in-sample and out-of-sample tests provide evidence that distress is captured more successfully with systematic variables and their interaction effects than with industry effects (which help only marginally). In unreported results, we also 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.

The importance of systematic variables in distress prediction is also demonstrated in Fig. 1, where we plot the predicted and observed distress rate for the period 2000–2009. The predicted distress rate is the simple average of the probabilities of distress of all firms in each period. The columns denote recession periods in the Eurozone as indicated by the OECD. The graph shows that in model I, where only firm-specific variables are included, the predicted distress rates follow a smooth upward trend, but do not co-vary with the observed distress rates. 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, predicted distress rates move together with observed ones and vary greatly with the business cycle, increasing with downturns and lowering with upturns. Once again, industry effects do not seem to provide additional improvements. These findings are in accordance with previous literature (Carling et al., 2007; Jacobson et al., 2005, 2013; Laerkholm-Jensen et al., 2015).

4.2. Regional models and generic unrestricted and restricted models

Now, we turn our analysis to the regional sub-samples presented in Section 3. First, we fit idiosyncratic and systematic variables for the regional sub-samples and estimate three regional models.[7] Second, we use the generic specifications of Section 4.1 and estimate three unrestricted generic models for the regional sub-samples. Our preferred model is model II of Table 5, because it considers both idiosyncratic and systematic variables, performs very well and, at the same time, has a simple specification. Third, we estimate a model for the overall sample that contains all the variables (not just a subset of them) and then impose the parameters for the systematic variables from this model in all regional sub-samples in order to estimate three restricted generic models. We would like to note that we allow the constant and the firm-specific parameters to vary in the regional sub-samples in order to ensure well-fitting restricted models in the cross-section. Lastly, we compare the generic unrestricted and restricted models with the regional models.

Table 6 presents the results. In Panels A–C, we present the results of the generic unrestricted models and the regional models. In Panels D–F, we present the results of the generic restricted models and repeat the results of the regional models to facilitate comparisons. In accordance with our hypothesis, we document performance improvements when we switch to the regional models. Interestingly, we find that the firm-specific variables identified as the most important in predicting distress in the regional models are exactly the same as in the generic unrestricted and restricted models. This is evidence that SMEs across Europe are sensitive to the same idiosyncratic factors. This does not hold in the case of systematic factors. Specifically, we document regional variations in the vulnerabilities to systematic factors, according to region-specific conditions and characteristics. Moreover, we do not find that the years taken to resolve insolvency variable adds predictive power in the regional models. This can partly be due to regional groups being relatively homogeneous with respect to their insolvency regimes.

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. When we move from the generic unrestricted to the regional model (Panel A), we document small changes in the coefficient sizes. Moving from the generic restricted to the regional model (Panel D), these changes are often even higher in magnitude. In the regional model we find that we need only two systematic variables to predict distress satisfactorily. These variables are the bank lending and the GDP growth. Both bank lending and GDP growth have significant coefficients and are, as expected, negatively related to the distress rate. Lower GDP growth means lower growth in sales by firms and thus an increased distress probability. Interestingly, even with less systematic variables, the regional model achieves higher performance than the generic ones. In Panel B, the percentage of distressed firms in the last three deciles increases from the generic unrestricted to the regional model (78.14–79.12%) and the AUC increases (0.806–0.825). In Panel E, these figures also increase when we move from the generic restricted to the regional model. Out-of-sample performance improvements (Panels C and F) are similar as in the case of in-sample results. The above provides evidence that SMEs in the countries of group 1, which consists of some of the strongest EU economies, are less sensitive to the macroeconomic situation. This is in accordance with the finding in Section 4.1, that large SMEs are less vulnerable to the macro-economic situation compared to small SMEs, since SMEs in group 1 are, on average, larger.

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. When we move from the generic unrestricted to the regional model (Panel A), we identify almost the same systematic factors as being the most useful for predicting distress. The years to resolve insolvency variable is replaced by the balance of payments variable, since these countries often suffer from current account deficits. Moving from the generic restricted to the regional model (Panel D), we find large changes in the coefficient sizes, especially for the systematic variables. 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 less favorable economic climate in the economies of group 2 during the years of this study. In Panel B, the regional model outperforms the generic unrestricted one (a 1.17% improvement in the correct classification for distressed firms and a 0.07 improvement in AUC). Similarly, in Panel E, the regional model outperforms the generic restricted one (a 1.78% improvement in the correct classification for distressed firms and a 0.02 modest improvement in AUC).

For group 3 (Czech Republic, Poland), the models are estimated from a sample of 48,470 SMEs (178,618 firm-year observations), which include 4278 distressed SMEs. When we move from the generic to the regional unrestricted model (Panel A), coefficient sizes of the idiosyncratic variables differ slightly and we observe an interesting new set of systematic variables. Moving from the generic to the regional restricted model (Panel D), both coefficient sizes of the idiosyncratic and systematic variables differ, the first ones slightly and the second ones quite substantially. In the regional model we find the FX volatility, the 10-year government bond yield and the GDP growth variables 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 and thus creates uncertainty about future economic conditions. Concerning the 10-year government bond yield variable, it enters with a positive coeffi-

[7] We also fit and estimate country models. Altman et al. (2014) show that the classification accuracy of the Z″-score (that uses only accounting data) can be considerably improved with country specific estimation. Findings are similar for countries of the same group, but country sub-samples often suffer from small size bias. Thus, in the sake of brevity and efficiency, we stick to regions instead of countries.

Table 6
Generic and regional models estimated for the regional subsamples. The models are estimated for 2000–2009 data with lagged yearly observations using the multi-period logit technique. The data-set is limited to non-financial SMEs. Group 1 has 165,786 French, German and British SMEs (801,536 firm-year observations) out of which 14,177 distressed. Group 2 has 429,978 Italian, Portuguese and Spanish SMEs (1,741,707 firm-year observations) out of which 30,900 distressed. Group 3 has 48,470 Czech and Polish SMEs (178,618 firm-year observations) out of which 4278 distressed. The regional models are estimated after fitting idiosyncratic and systematic variables to the regional sub-samples. The generic unrestricted models are estimated using the generic model II of Table 5 and fitting it to the regional sub-samples. The generic restricted models are estimated using all the idiosyncratic and systematic variables that are identified in the generic and regional specifications and imposing the same parameters to the systematic variables for the regional sub-samples (whereas the parameters of the idiosyncratic variables are allowed to differ). Standard errors are firm clustered-corrected (Huber/White standard errors). Parameter estimates are given first followed by chi-square values in parentheses. * denotes significance at a 5% level, ** at a 1% level and *** at a 0.1% level. FX rate is calculated in relation to the US dollar (USD/national currency). FX rate volatility values are rolling annual averages at daily frequency. Unemployment is measured in%. The economic sentiment indicator is calculated by the Directorate General of Financial Affairs of the European Commission as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program.

	Group 1				Group 2				Group 3			
	Generic unrestricted model		Regional model		Generic unrestricted model		Regional model		Generic unrestricted model		Regional model	
Panel A. Estimation results, generic unrestricted models versus regional models												
Earnings before taxes to total assets	−1.0667***	(−15.97)	−1.077***	(−15.54)	−0.681***	(−10.67)	−0.677***	(−10.61)	−0.534***	(−4.50)	−0.547***	(−4.68)
EBITDA to interest expenses	−0.0000470***	(−11.10)	−0.0000483***	(−11.31)	−0.0000472***	(−9.80)	−0.0000468***	(−9.82)	−0.0000507***	(−4.62)	−0.0000509***	(−4.55)
Current liabilities to total assets	1.908***	(71.61)	1.916***	(72.99)	1.216***	(69.49)	1.217***	(69.66)	1.415***	(33.07)	1.397***	(32.84)
Cash flow to current liabilities	−0.236***	(−4.15)	−0.196**	(−3.16)	−0.654***	(−9.51)	−0.650***	(−9.43)	−0.336***	(−2.75)	−0.314**	(−2.64)
Turnover to total liabilities	−0.106***	(−14.72)	−0.101***	(−14.07)	−0.245***	(−30.10)	−0.249***	(−30.41)	−0.171***	(−13.05)	−0.175***	(−13.28)
Size (ln(totals assets))	−0.0226***	(−3.05)	−0.00559	(−0.75)	−0.101***	(−16.95)	−0.107***	(−17.72)	−0.0587***	(−4.60)	−0.0754***	(−6.04)
Dummy equal to 1 if shareholders are more than 2	−0.0781***	(−3.44)	−0.0812***	(−3.62)	−0.334***	(−20.51)	−0.324***	(−19.96)	−0.344***	(−7.49)	−0.347***	(−7.53)
Dummy equal to 1 if SME is located in urban area	0.151***	(4.82)	0.174***	(5.68)	0.103***	(6.50)	0.101***	(6.43)	0.351***	(9.75)	0.358***	(9.92)
Loans granted to non-financial sector (% change)	−2.388***	(−14.95)	−4.611***	(−25.53)	−0.268	(−1.29)	−3.378***	(−30.14)	−6.203***	(−21.64)		
Years to resolve insolvency proceedings	−1.206***	(−21.27)			1.171***	(18.92)			−0.237***	(−20.25)		
GDP growth (% change)			−5.595***	(−9.44)							−11.62***	(−22.52)
FX rate (% change)	−2052.2***	(−49.78)			−2403.5***	(−46.89)	−2276.6***	(−44.99)	340.15***	(3.06)		
Unemployment	19.425***	(13.59)			13.441***	(27.02)	6.176***	(24.91)	−22.213***	(−15.94)		
Economic sentiment	−0.0206***	(−15.16)			−0.0279***	(−23.62)	−0.0256***	(−21.08)	−0.0031	(−0.93)		
Balance of Payments (% GDP)							−42.082***	(−16.28)				
FX rate volatility											122.6***	(12.11)
10-year government bond yield											25.43***	(14.63)
Constant	Yes		Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes		Yes	
Firm-year observations	801,536		801,536		1,741,707		1,741,707		178,618		178,618	
Firms	165,786		165,786		429,978		429,978		48,470		48,470	
Distressed firms	14,177		14,177		30,900		30,900		4278		4278	
Pseudo R-squared	0.150		0.150		0.170		0.177		0.214		0.250	
Log likelihood	−61,573.50		−60,538.70		−131,451.70		−127,673.50		−15,878.40		−15,147.90	
Wald test	19,302.49***		20,225.9***		55,513.51***		55,783.8***		8099.98***		8083.9***	
Panel B. In-sample prediction tests, generic unrestricted models versus regional models												
Hosmer–Lemeshow decile												
1–5	9.24%		10.17%		9.47%		9.84%		6.83%		6.77%	
8	9.95%		10.07%		10.76%		9.96%		10.34%		10.43%	
9	15.29%		16.69%		15.31%		17.34%		18.94%		19.23%	
10	52.90%		52.36%		52.21%		52.15%		54.17%		53.95%	
8–10	78.14%		79.12%		78.29%		79.46%		83.44%		83.61%	
Area under the ROC curve	0.806		0.825		0.841		0.848		0.853		0.875	
Panel C. Out-of-sample prediction tests, generic unrestricted models versus regional models												
For Group 1, a hold-out sample of 18,449 French, German and British SMEs (88,957 firm-year observations) is used. For Group 2, a hold-out sample of 48,034 Italian, Portuguese and Spanish SMEs (195,236 firm-year observations) is used. For Group 3, a hold-out sample of 5340 Czech and Polish SMEs (19,844 firm-year observations) is used												
Hosmer–Lemeshow decile												
1–5	6.84%		6.73%		6.52%		5.98%		5.62%		6.09%	
8	11.07%		11.12%		7.18%		8.42%		6.56%		6.09%	

Table 6 (*continued*)

	Group 1				Group 2				Group 3			
	Generic unrestricted model		Regional model		Generic unrestricted model		Regional model		Generic unrestricted model		Regional model	
9	18.35%		18.64%		15.97%		14.91%		18.03%		18.03%	
10	54.37%		54.25%		62.79%		63.28%		59.72%		60.89%	
8–10	83.78%		84.01%		85.93%		86.61%		84.31%		85.01%	
Area under the ROC curve	0.805		0.824		0.843		0.850		0.841		0.868	
Panel D. Estimation results, generic restricted models versus regional models												
Earnings before taxes to total assets	−1.163***	(−16.81)	−1.077***	(−15.54)	−0.705***	(11.09)	−0.677***	(−10.61)	−0.573***	(−4.84)	−0.547***	(−4.68)
EBITDA to interest expenses	−0.0000453***	(−10.44)	−0.0000483***	(−11.31)	−0.0000499***	(−10.06)	−0.0000468***	(−9.82)	−0.0000482***	(−4.10)	−0.0000509***	(−4.55)
Current liabilities to total assets	1.959***	(72.55)	1.916***	(72.99)	1.208***	(69.09)	1.217***	(69.66)	1.380***	(31.59)	1.397***	(32.84)
Cash flow to current liabilities	−0.154**	(−2.48)	−0.196**	(−3.16)	−0.613***	(−8.93)	−0.650***	(−9.43)	−0.292***	(−2.41)	−0.314**	(−2.64)
Turnover to total liabilities	−0.102***	(13.80)	−0.101***	(−14.07)	−0.251***	(−31.02)	−0.249***	(−30.41)	−0.185***	(−13.60)	−0.175***	(−13.28)
Size (ln(totals assets))	−0.02242**	(−2.94)	−0.00559	(−0.75)	−0.141***	(−26.00)	−0.107***	(−17.72)	−0.113***	(−9.59)	−0.0754***	(−6.04)
Dummy equal to 1 if shareholders are more than 2	−0.0463*	(−2.06)	−0.0812***	(−3.62)	−0.386***	(−24.02)	−0.324***	(−19.96)	−0.335***	(−7.13)	−0.347***	(−7.53)
Dummy equal to 1 if SME is located in urban area	0.215***	(7.00)	0.174***	(5.68)	0.088***	(5.57)	0.101***	(6.43)	−0.366***	(9.96)	0.358***	(9.92)
Loans granted to non-financial sector (% change)	−3.533		−4.611***	(−25.53)	−3.533		−3.378***	(−30.14)	−3.533			
Years to resolve insolvency proceedings	0.077				0.077				0.077			
GDP growth (% change)	−2.570		−5.595***	(−9.44)	−2.570				−2.570		−11.62***	(−22.52)
FX rate (% change)	−1757.6				−1757.6		−2276.6***	(−44.99)	−1757.6			
Unemployment	2.260				2.260		6.176***	(24.91)	2.260			
Economic sentiment	−0.0221				−0.0221		−0.0256***	(−21.08)	−0.0221			
Balance of Payments (% GDP)	0.3953				0.3953		−42.082***	(−16.28)	0.3953			
FX rate volatility	49.47				49.47				49.47		122.6***	(12.11)
10-year government bond yield	8.13				8.13				8.13		25.43***	(14.63)
Constant	Yes		Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes		Yes	
Firm-year observations	801,536		801,536		1,741,707		1,741,707		178,618		178,618	
Firms	165,786		165,786		429,978		429,978		48,470		48,470	
Distressed firms	14,177		14,177		30,900		30,900		4278		4278	
Log likelihood	−59,507.75		−60,538.70		−128,109.14		−127,673.50		−15,823.64		−15,147.90	
Wald test	18,444.02***		20,225.9***		46,580.3***		55,783.8***		7743.0***		8083.9***	
Panel E. In-sample prediction tests, generic restricted models versus regional models												
Hosmer–Lemeshow decile												
1–5	10.33%		10.17%		10.21%		9.84%		7.37%		6.77%	
8	10.06%		10.07%		9.53%		9.96%		9.89%		10.43%	
9	15.02%		16.69%		16.42%		17.34%		18.92%		19.23%	
10	52.39%		52.36%		51.72%		52.15%		54.33%		53.95%	
8–10	77.46%		79.12%		77.68%		79.46%		83.14%		83.61%	
Area under the ROC curve	0.825		0.825		0.846		0.848		0.856		0.875	
Panel F. Out-of-sample prediction tests, generic restricted models versus regional models												
For Group 1, a hold-out sample of 18,449 French, German and British SMEs (88,957 firm-year observations) is used. For Group 2, a hold-out sample of 48,034 Italian, Portuguese and Spanish SMEs (195,236 firm-year observations) is used. For Group 3, a hold-out sample of 5340 Czech and Polish SMEs (19,844 firm-year observations) is used												
Hosmer–Lemeshow decile												
1–5	7.37%		6.73%		8.09%		5.98%		9.84%		6.09%	
8	10.02%		11.12%		8.18%		8.42%		7.73%		6.09%	
9	18.29%		18.64%		17.51%		14.91%		16.86%		18.03%	
10	54.75%		54.25%		57.57%		63.28%		57.14%		60.89%	
8–10	83.05%		84.01%		83.26%		86.61%		81.73%		85.01%	
Area under the ROC curve	0.823		0.824		0.847		0.850		0.840		0.868	

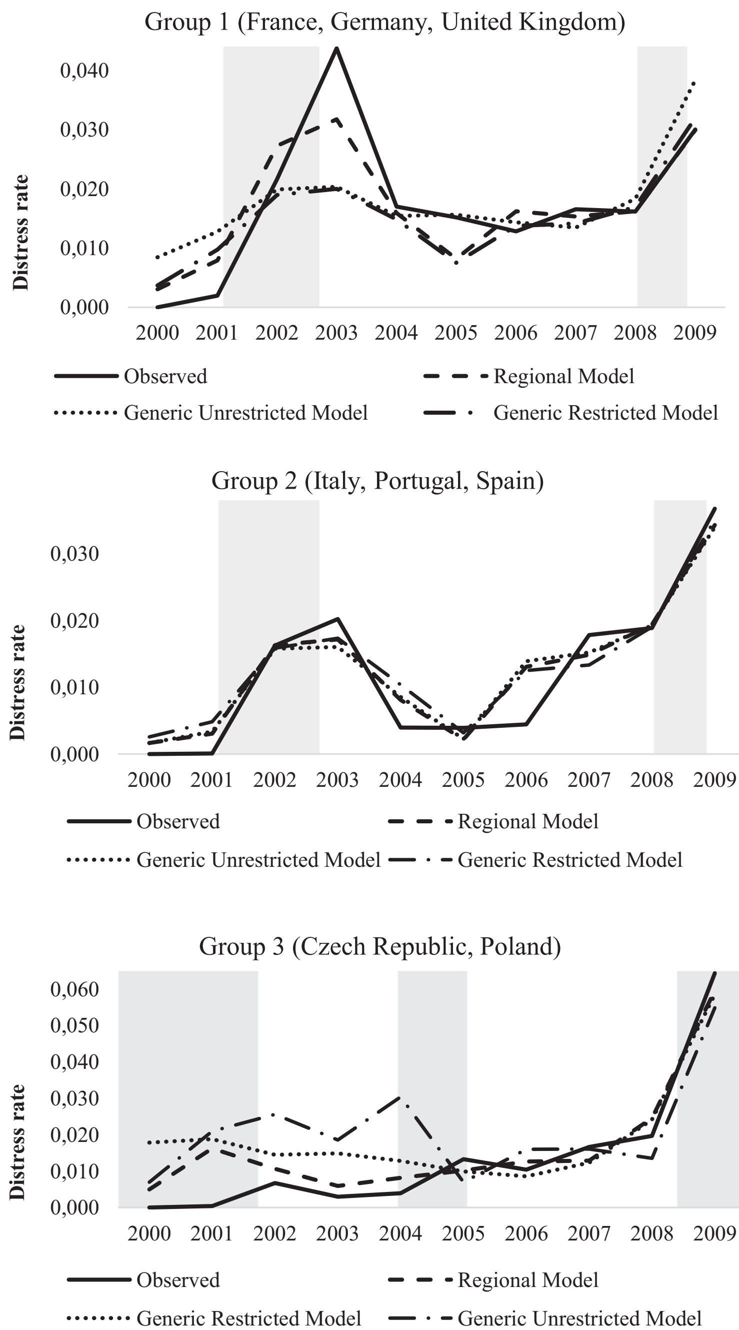

Fig. 2. Predicted and observed distress rates for the 3 groups. The figure plots the predicted distress rates based on the regional and generic models of Table 6, along with the observed distress rate for each group. The predicted distress rate is the simple average of the probabilities of distress of all firms in each group and period. The regional models are estimated after fitting idiosyncratic and systematic variables to the regional sub-samples. The generic unrestricted models are estimated using the generic model II of Table 5 and fitting it to the regional sub-samples. The generic restricted models are estimated using all the idiosyncratic and systematic variables that are identified in the generic and regional specifications and imposing the same parameters to the systematic variables for the regional sub-samples (whereas the parameters of the idiosyncratic variables are allowed to differ). For groups 1 and 2, the columns denote recession periods in the euro area, and for group 3, recession periods in the Czech Republic and Poland, as indicated by OECD.

cient. 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.

According to the Hosmer–Lemeshow grouping, the percentage of distressed firms in the last three deciles increases slightly from the generic unrestricted to the regional model (Panel B) and from the generic restricted to the regional model (Panel D). The increase in the AUC is more obvious (0.853–0.875 in Panel B and 0.856–0.875 in Panel D). Clearly, the specific set of systematic variables helps in better capturing distress risk. The out-of-sample results (Panels C and F) give the same picture.

The above evidence shows that the fit is improved when we change some of the macroeconomic covariates as we move from the generic (unrestricted and restricted) to the regional models. This indicates that regional models are better able to capture the

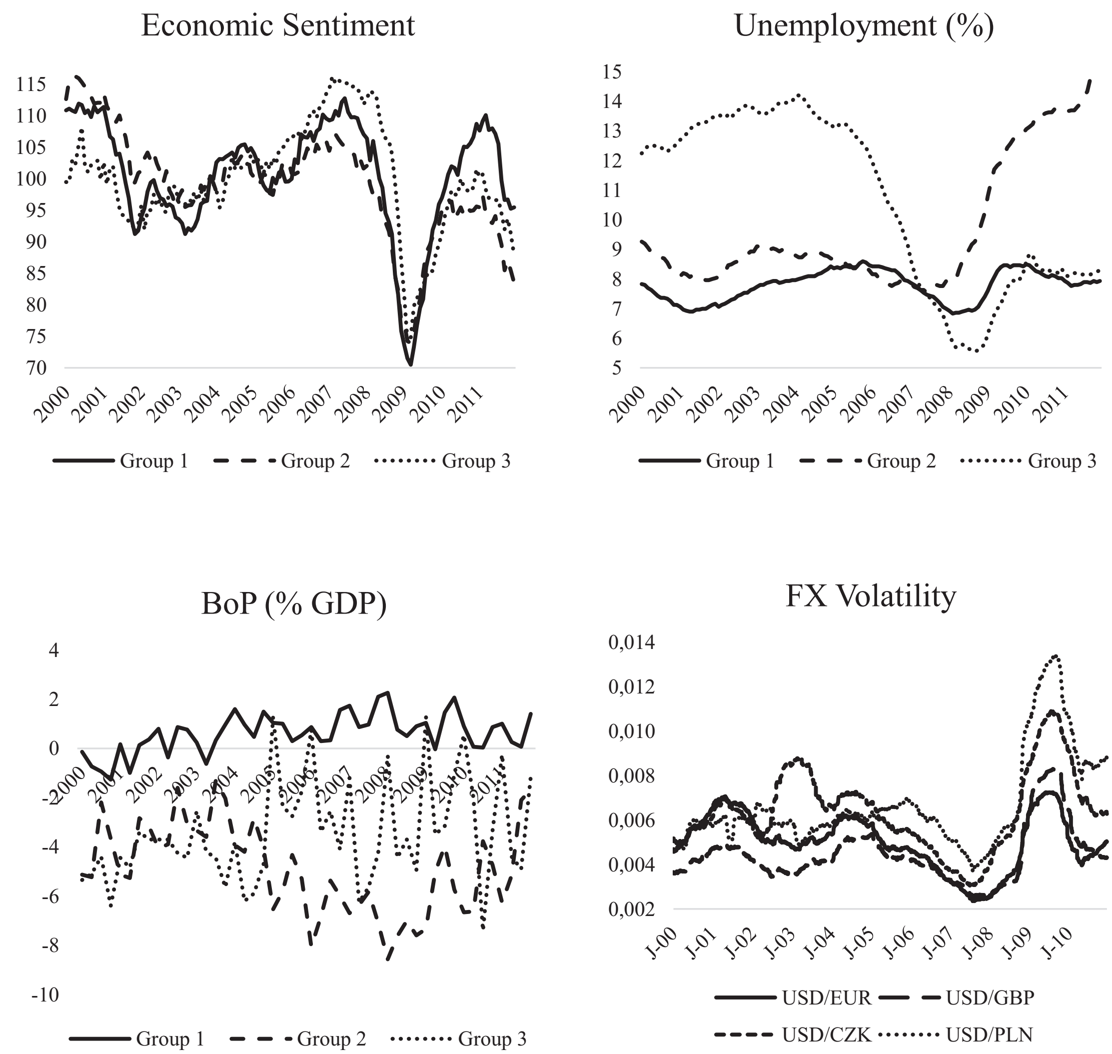

Fig. 3. Macroeconomic variables 2000–2011. The figure plots the aggregate time series for four macroeconomic variables. The economic sentiment indicator and unemployment values are rolling annual averages at monthly frequency. Balance of payments values are rolling annual averages at quarterly frequency. Foreign exchange rate volatility values are rolling annual averages at daily frequency.

systematic effects. Although the improvements might seem small, Figs. 2 and 3 give a better sense of the comparative performance.

Fig. 2 plots the predicted distress rate based on the regional and generic unrestricted and restricted models of Table 6, along with the observed distress rate for each group. It is obvious that the predicted distress rates from the three regional models match better the observed values, compared to the predicted distress rate from the generic unrestricted and restricted models. For group 1, both the generic unrestricted and restricted models underestimate the observed distress rate for the early years of our study (before 2004). For 2008 and 2009, the generic unrestricted model overestimates the observed distress rate. The generic restricted model moves closely together with the regional model. In the case of group 2, the three time-series are very similar. This is probably due to the vast majority of companies in our sample belonging to group 2, thus the regional and generic models for this group give almost identical predictions. Finally, for group 3, the generic models overestimate the observed distress rates for the years 2000–2005 more than the regional model does and co-move together with the regional model (and the observed values) in later years.

In unreported results, we use the estimated coefficients from 2000–2009 to predict the distress rates for 2000–2011 and then we compare these predicted distress rates with the observed ones for 2010–2011. We use as observed distressed rates for 2010–11 those documented by Creditreform (Creditreform, 2012) which is the largest private credit bureau in Germany that gathers statistics on insolvencies in Europe. Unfortunately, our models do not always perform well out-of-sample. Specifically, there are significant out-of-sample forecast errors for group 2. To check whether this is due to omission of systematic variables, we estimate several models where we keep all the firm-specific variables of Table 6 but we try out different systematic variables. Out-of-sample performance does not improve though, thus we conclude that the poor out-of-sample performance for group 2 from 2010 onwards is not due to omission of variables.

Fig. 3 plots the aggregate time series for four macroeconomic variables. They are economic sentiment, unemployment (in percentage terms), the balance of payments (as a percentage of the GDP), and foreign exchange rate volatility. The economic sentiment, the unemployment and the balance of payments variables are average values for the countries in each group. The foreign exchange rate is in relation to the US dollar. We report volatility for each currency separately (euro, British pound, Czech koruna and Polish zloty).

Table 7

Robustness test with different distress definitions. The models are estimated for 2000–2009 data with lagged yearly observations using the multi-period logit technique. The data-set is limited to non-financial SMEs. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain, and group 3 includes Czech Republic and Poland. According to the main distress definition, a firm-year is distressed if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) appears with one of the statuses "defaulted", "in receivership", "bankrupt", "in liquidation" or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. In the alternative distress definition 1, we exclude all firms that disappear from the sample before 2010 without updated status situation. These include firms that, under the main distress definition are classified as distressed if their equity is negative in the last year. Thus, the alternative distress definition 1 is strictly linked to a legal insolvency procedure. In the alternative distress definition 2, we exclude all firms that have negative equity in one or more of the years of their existence in the sample. Thus, under the alternative distress definition 2, we essentially lose some distress-related information because we include in the sample only firms with non-negative equity. Standard errors are firm clustered-corrected (Huber/White standard errors). Parameter estimates are given first followed by chi-square values in parentheses. * denotes significance at a 5% level, ** at a 1% level and *** at a 0.1% level. FX rate is calculated in relation to the US dollar (USD/national currency). FX rate volatility values are rolling annual averages at daily frequency. Unemployment is measured in%. The economic sentiment indicator is calculated by the Directorate General of Financial Affairs of the European Commission as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program.

	Overall sample – generic model						Group 1 – regional model					
	Main definition		Alternative definition 1		Alternative definition 2		Main definition		Alternative definition 1		Alternative definition 2	
Earnings before taxes to total assets	−0.770***	(−15.89)	−0.699***	(−5.43)	−3.977***	(−23.86)	−1.077***	(−15.54)	−0.525***	(−3.36)	−3.063***	(−13.10)
EBITDA to interest expenses	−0.0000450***	(−14.49)	−0.0000289***	(−6.33)	−0.0000311***	(−5.08)	−0.0000483***	(−11.31)	−0.0000333***	(−6.15)	−0.0000252***	(−3.84)
Current liabilities to total assets	1.420***	(103.81)	1.536***	(49.76)	1.860***	(27.45)	1.916***	(72.99)	2.007***	(44.26)	2.771***	(23.24)
Cash flow to current liabilities	−0.485***	(−9.14)	−1.277***	(−11.35)	−0.143*	(−2.11)	−0.196**	(−3.16)	−1.149***	(−8.13)	−0.460***	(−3.20)
Turnover to total liabilities	−0.177***	(−36.08)	−0.035***	(−5.67)	−0.003	(−0.49)	−0.101***	(−14.07)	−0.017*	(−2.17)	0.056	(1.74)
Size (ln(totals assets))	−0.0940***	(−22.95)	0.376***	(45.80)	0.233***	(23.92)	−0.00559	(−0.75)	0.308***	(27.33)	0.226***	(16.61)
Dummy equal to 1 if shareholders are more than 2	−0.274***	(−21.99)	−0.214***	(−9.48)	−0.203***	(−6.99)	−0.0812***	(−3.62)	−0.0585	(−1.79)	−0.0338	(−0.78)
Dummy equal to 1 if SME is located in an urban area	0.141***	(10.85)	0.0341	(1.15)	0.0147	(0.40)	0.174***	(5.68)	0.239***	(5.05)	0.444***	(7.99)
FX rate (% change)	−1686.8	(−59.04)	−1357.1***	(−27.83)	−1100.8***	(−19.41)						
Unemployment	1.883***	(12.39)	1.502***	(28.97)	1.317***	(9.11)						
Economic sentiment indicator	−0.0259***	(−35.03)	−0.0047***	(−3.22)	−0.0022	(1.26)						
Loans granted to non-financial sector (% change)	−4.414***	(−58.29)	−5.194***	(−34.27)	−3.532***	(−21.30)	−4.611***	(−25.53)	−6.404***	(−20.77)	−2.413***	(−5.63)
Years to resolve insolvency proceedings	0.0949***	(27.40)	0.732***	(21.49)	0.0087	(0.85)						
GDP growth (% change)							−5.595***	(−9.44)	0.963	(0.00)	−2.976***	(−2.37)
Constant	Yes		Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes		Yes	
Firm-year observations	2,721,861		1,594,433		2,245,724		801,536		332,547		686,976	
Firms	644,234		389,347		520,636		165,786		66,306		140,196	
Distressed firms	49,355		12,362		6836		14,177		5646		3073	
Pseudo R-squared	0.171		0.115		0.072		0.150		0.098		0.062	
Log likelihood	−204,638.5		−60,050.1		−43,097.6		−60,538.7		−25,683.6		−18,477.7	
Wald test	84,259.5***		16,563.99***		7117.25***		20,225.9***		5359.19***		2485.26***	
Area under the ROC curve	0.838		0.794		0.769		0.825		0.776		0.746	

	Group 2 – regional model						Group 3 – regional model					
	Main definition		Alternative definition 1		Alternative definition 2		Main definition		Alternative definition 1		Alternative definition 2	
Earnings before taxes to total assets	−0.677***	(−10.61)	−1.248***	(−6.63)	−4.746***	(−18.26)	−0.547***	(−4.68)	−3.092***	(−4.45)	−4.032**	(−4.64)
EBITDA to interest expenses	−0.0000468***	(−9.82)	−0.0000507***	(−5.21)	−0.0000823***	(−4.02)	−0.0000509***	(−4.55)	0.0000009	(0.03)	0.00000212	(0.09)
Current liabilities to total assets	1.217***	(69.66)	1.126***	(26.39)	1.568***	(18.75)	1.397***	(32.84)	1.204***	(3.21)	2.117***	(5.91)
Cash flow to current liabilities	−0.650***	(−9.43)	−1.111***	(−7.27)	−0.099	(−1.06)	−0.314**	(−2.64)	−0.149***	(−0.68)	−0.0301	(−0.26)
Turnover to total liabilities	−0.249***	(−30.41)	−0.0925***	(−7.25)	−0.0415***	(−3.38)	−0.175***	(−13.28)	−0.0195	(−0.86)	−0.0707***	(−5.38)
Size (ln(totals assets))	−0.107***	(−17.72)	0.462***	(35.44)	0.474***	(29.68)	−0.0754***	(−6.04)	−0.1228	(−1.16)	0.345***	(5.83)
Dummy equal to 1 if shareholders are more than 2	−0.324***	(−19.96)	−0.368***	(−11.36)	−0.344***	(−8.38)	−0.347***	(−7.53)	−0.229	(−1.15)	−0.504**	(−2.61)
Dummy equal to 1 if SME is located in urban area	0.101***	(6.43)	0.0037	(0.10)	0.0509	(0.99)	0.358***	(9.92)	0.354	(1.20)	0.119	(0.65)
Loans granted to non-financial sector (% change)	−3.378***	(−30.14)	−4.482***	(−20.47)	−3.987***	(−15.94)						
GDP growth (% change)							−11.62***	(−22.52)	0.700	(0.11)	−9.504*	(−2.07)
FX rate (% change)	−2276.6***	(−44.99)	−808.86***	(−11.98)	−681.11***	(−6.81)						
Unemployment	6.176***	(24.91)	6.709***	(9.16)	6.357***	(3.56)						

Table 7 (*continued*)

	Overall sample – generic model						Group 1 – regional model					
	Main definition		Alternative definition 1		Alternative definition 2		Main definition		Alternative definition 1		Alternative definition 2	
Economic sentiment	−0.0256***	(−21.08)	−0.01888***	(−6.56)	−0.04122***	(−11.77)						
Balance of payments (% GDP)	−42.082***	(−16.28)	−30.575***	(−35.61)	−39.704***	(−46.14)						
FX rate volatility							122.6***	(12.11)	308.32***	(4.85)	257.18***	(5.18)
10-year government bond yield							25.43***	(14.63)	57.14*	(2.85)	4.031	(0.51)
Constant	Yes		Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes		Yes	
Firm-year observations	1,741,707		1,185,258		1,412,358		178,618		76,628		146,390	
Firms	429,978		302,959		341,471		48,470		20,082		38,969	
Distressed firms	30,900		6338		3566		4278		378		197	
Pseudo R-squared	0.177		0.127		0.138		0.250		0.214		0.103	
Log likelihood	−127,673.5		−32,418.6		−21,454.16		−15,147.9		−1247.1		−1345.6	
Wald test	55,783.8***		9052.84***		6654.5***		8083.9***		611.51***		320.46***	
Area under the ROC curve	0.848		0.847		0.854		0.879		0.816		0.820	

The economic sentiment indicator clearly captures the deep recession in 2008–2009. We find this variable to be an important distress determinant in the generic model for the overall sample and also in the regional model for group 2. We can see that from 2006 onwards, values of the economic sentiment indicator for group 2 are systematically lower than for groups 1 and 3, capturing the higher sensitivity of the distress rate for group 2 to the values of this indicator.

The same situation holds for unemployment as well. Specifically, unemployment is relatively stable during the years of the study for group 1. For groups 2, it is increasing substantially from the economic slowdown of 2008 onwards. Group 3 experiences a substantial decrease for 2004–2009 and a moderate increase later. Not surprisingly, we find unemployment to be an important distress determinant in the regional model for group 2.

The balance of payments as a percentage of GDP is also relatively stable (values around zero) during the years of this study for group 1. For groups 2, values are always negative and usually much lower than for group 3. Again, not surprisingly, it has a significant impact in the regional model for group 2, but not for the regional models of the other groups. This is evidence that SME distress rates in the countries of group 2 are particularly sensitive to the high current account deficits of their economies.

Finally, the volatility of foreign exchange rates against the dollar follows similar trends for all currencies, but it is the Czech koruna and Polish zloty that have the highest volatility values. Thus, in the regional model for group 3, we identify this variable as a significant determinant of the SME distress rate.

5. Robustness tests

5.1. Definition of distress

In Section 3, we discuss that in around 40% of our sample (254,887 out of 644,234 firms), a firm disappears from the database before 2010 but the status information remains outdated. We also extensively discuss the challenges in tracking properly the status of SMEs that lead us to adopt an assumption for this 40% of firms. Thus, under our main distress definition, in order to separate the cases of closure from the ones of distress, we assume that the last available firm-year for this 40% of firms is distressed if equity is negative in this last year. Because this assumption influences a large percentage of our sample, the estimation results can be sensitive to it. Therefore, in this section, we perform two robustness tests, in each one applying an alternative distress definition.

In the first alternative distress definition we exclude the 254,887 firms (1,127,428 firm-years) that disappear from the sample before 2010 without updated status situation. As described above, these firms correspond to around 40% of our sample and include all those that, under the main distress definition, are classified as distressed if their equity is negative in the last year. We would like to clarify here that these firms do not necessarily have negative equity in the last year. Specifically, under the main distress definition, firms that disappear from the sample before 2010 without updated status situation are classified as healthy if their equity is non-negative in the last year. By excluding all firms that disappear from the sample before 2010 without updated status situation, regardless if they have negative equity or not, we actually exclude all firms for which we use the negative equity condition in order to classify them as distressed or healthy (simply because we lack information on their actual status situation). So, under the first alternative distress definition, distress is strictly linked to only a legal insolvency procedure. In particular, here the last available firm-year of firms that appear with one of the fol-

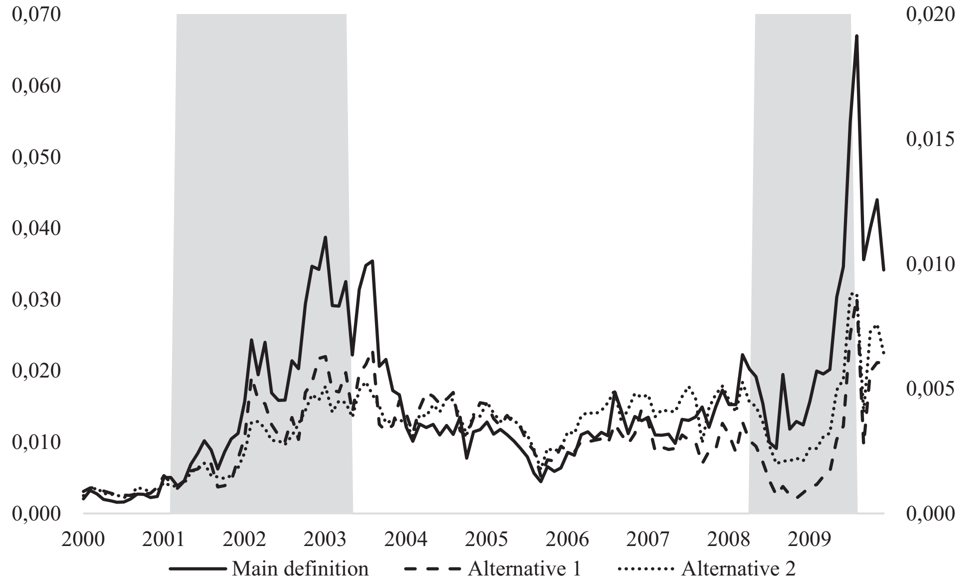

Fig. 4. Predicted distress rates with different distress definitions. The figure plots the predicted distress rate based on the generic modes of Table 7 for the overall sample under the main distress definition (left scale), the alternative distress definition 1 (left scale) and the alternative distress definition 2 (right scale). According to the main distress definition, a firm-year is distressed if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) appears with one of the statuses "defaulted", "in receivership", "bankrupt", "in liquidation" or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. In the alternative distress definition 1, we exclude all firms that disappear from the sample before 2010 without updated status situation. These include firms that, under the main distress definition are classified as distressed if their equity is negative in the last year. Thus, the alternative distress definition 1 is strictly linked to a legal insolvency procedure. In the alternative distress definition 2, we exclude all firms that have negative equity in one or more of the years of their existence in the sample. Thus, under the alternative distress definition 2, we essentially lose some distress-related information because we include in the sample only firms with non-negative equity. The columns denote recession periods in the Euro area, as indicated by OECD.

lowing statuses "defaulted", "in receivership", "bankrupt", "in liquidation" is distressed, and firm-years of firms that appear as "active" during the whole sample period (i.e. up to 2010) are healthy. Thus, we include in the sample only firms that we know with certainty whether they are distressed or healthy. The sample includes 1,594,433 firm-years (389,374 firms) out of which 12,362 are distressed.

In the second alternative distress definition, we exclude the 123,688 firms (476,137 firm-years) that have negative equity in one or more of the years of their existence in the sample. These firms correspond to around 20% of our sample and include, but are not limited to, all the firms that, under the main distress definition, are classified as distressed because their equity is negative in the last year. We would like to clarify here that, under the second alternative distress definition, we also exclude firms that have negative equity but are either not classified as distressed under the main distress definition (because they appear with the status situation "active" during the whole sample period) or are classified as distressed under the main distress definition but the primary reason for this is not the fact that they have negative equity but the fact that they appear as "defaulted" or "in receivership" or "bankrupt" or "in liquidation". By excluding all firms that have negative equity at some point of their existence in the sample, we set a sample selection rule and apply it uniformly. So, under the second alternative distress definition, we essentially lose some distress-related information because we include in the sample only firms with non-negative equity for all their firm-years. The sample includes 2,245,724 firm-years (520,636 firms) out of which 6836 are distressed.

Appendix C reports comparative statistics for our two alternative distress definitions versus our main distress definition. It is obvious from Appendix C that our sample size decreases significantly and distress rates are much lower in both cases. Nevertheless, as we describe below, our results remain robust.

Table 7 reports the estimation results for the overall sample and for the regional groups under our main distress definition and the two alternative distress definitions for comparison purposes. Almost all variables retain their signs and significance. In a few cases that the sign flips, coefficients do not demonstrate significance. The exception is size, which, in the vast majority of cases, has a significantly negative coefficient under the main distress definition and a significantly positive coefficient under both alternative distress definitions. An explanation for this can be its non-linear effect. Specifically, as mentioned before, we find a positive coefficient for squared size, indicating that for very large SMEs, distress risk starts to increase, probably because these companies are more likely to be pursued by their creditors in liquidation procedures. Further supporting this argument, we find that the 12,362 distressed cases under the first alternative distress definition come from 2 times larger companies and the 6836 distressed cases under the second alternative distress definition come from 1.5 times larger companies compared to the 49,355 distressed cases under the main definition.

Regarding the performance of the models, we report the pseudo R^2 and AUC. We find them to be always higher under the main distress definition than under the two alternative definitions.

To further verify our point, Fig. 4 plots the predicted distress rate based on the generic modes of Table 7 for the overall sample under the main distress definition (left scale), the first alternative distress definition (left scale) and second alternative distress definition (right scale). The columns denote recession periods in the Euro area, as indicated by OECD. All measures display similar patterns and commove together. We can clearly see though that the main distress definition makes our estimates more sensitive to the business cycle.

5.2. Estimation technique

In addition to the multi-period logit model developed by Shumway (2001), we apply the Cox proportional hazard model (Cox, 1972) that makes different assumptions about the hazard

Table 8
Robustness test with different estimation techniques. The models are estimated for 2000–2009 data with lagged yearly observations using the multi-period logit and Cox proportional hazard techniques. The data-set is limited to non-financial SMEs. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain, and group 3 includes Czech Republic and Poland. The Cox proportional hazard model makes different assumptions about the hazard function. We follow Laerkholm-Jensen et al. (2015) and estimate a fully parametric model with a constant baseline intensity, since the usual Cox semi-parametric model does not allow us to simultaneously identify the vector of macroeconomic coefficients as well as the time-varying baseline intensity. Parameter estimates are given first followed by chi-square values in parentheses. * denotes significance at a 5% level, ** at a 1% level and *** at a 0.1% level. FX rate is calculated in relation to the US dollar (USD/national currency). FX rate volatility values are rolling annual averages at daily frequency. Unemployment is measured in%. The economic sentiment indicator is calculated by the Directorate General of Financial Affairs of the European Commission as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program.

	Overall sample - generic model				Group 1 - regional model			
	Logit		Cox		Logit		Cox	
Earnings before taxes to total assets	−0.770***	(−15.89)	−0.496***	(−11.30)	−1.077***	(−15.54)	−0.719***	(−10.21)
EBITDA to interest expenses	−0.0000450***	(−14.49)	−0.0000413***	(−13.79)	−0.0000483***	(−11.31)	−0.0000463***	(−11.11)
Current liabilities to total assets	1.420***	(103.81)	1.197***	(94.38)	1.916***	(72.99)	1.653***	(69.97)
Cash flow to current liabilities	−0.485***	(−9.14)	−0.667***	(−13.15)	−0.196**	(−3.16)	−0.367***	(−4.69)
Turnover to total liabilities	−0.177***	(−36.08)	−0.183***	(−37.28)	−0.101***	(−14.07)	−0.088***	(−12.91)
Size (ln(totals assets))	−0.0940***	(−22.95)	−0.084***	(−21.64)	−0.00559	(−0.75)	0.01882**	(2.95)
Dummy equal to 1 if shareholders are more than 2	−0.274***	(−21.99)	−0.239***	(−20.17)	−0.0812***	(−3.62)	−0.0635**	(−2.98)
Dummy equal to 1 if SME is located in an urban area	0.141***	(10.85)	0.096***	(7.95)	0.174***	(5.68)	0.122***	(4.24)
FX rate (% change)	−1686.8	(−59.04)	−1445.2***	(−52.53)				
Unemployment	1.883***	(12.39)	3.084***	(21.66)				
Economic sentiment indicator	−0.0259***	(−35.03)	−0.0066***	(−8.46)				
Loans granted to non-financial sector (% change)	−4.414***	(−58.29)	−3.624***	(−53.55)	−4.611***	(−25.53)	−2.713***	(−19.34)
Years to resolve insolvency proceedings	0.0949***	(27.40)	0.0993***	(30.72)				
GDP growth (% change)					−5.595***	(−9.44)	−3.844***	(−7.45)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	2,721,861		2,721,861		801,536		801,536	
Firms	644,234		644,234		165,786		165,786	
Distressed firms	49,355		49,355		14,177		14,177	
Log likelihood	−204,638.50		−91,145.41		−60,538.70		−25,643.69	
Wald test	84,259.5***		97,333.98***		20,225.9***		26,984.78***	
	Group 2 – regional model				**Group 3 – regional model**			
	Logit		Cox		Logit		Cox	
Earnings before taxes to total assets	−0.677***	(−10.61)	−0.421***	(−7.54)	−0.547***	(−4.68)	−0.304**	(−2.86)
EBITDA to interest expenses	−0.0000468***	(−9.82)	−0.0000413***	(−9.20)	−0.0000509***	(−4.55)	−0.0000451***	(−3.74)
Current liabilities to total assets	1.217***	(69.66)	1.026***	(63.38)	1.397***	(32.84)	1.135***	(27.85)
Cash flow to current liabilities	−0.650***	(−9.43)	−0.803***	(−13.08)	−0.314**	(−2.64)	−0.416**	(−3.26)
Turnover to total liabilities	−0.249***	(−30.41)	−0.239***	(−30.81)	−0.175***	(−13.28)	−0.171***	(−12.25)
Size (ln(totals assets))	−0.107***	(−17.72)	−0.097***	(−17.41)	−0.0754***	(−6.04)	−0.0355**	(−3.17)
Dummy equal to 1 if shareholders are more than 2	−0.324***	(−19.96)	−0.302***	(−19.71)	−0.347***	(−7.53)	−0.283***	(−6.28)
Dummy equal to 1 if SME is located in urban area	0.101***	(6.43)	0.079***	(5.47)	0.358***	(9.92)	0.276***	(8.10)
Loans granted to non-financial sector (% change)	−3.378***	(−30.14)	−4.178***	(−40.2)			−10.25***	(−17.42)
GDP growth (% change)					−11.62***	(−22.52)	−10.25***	(−17.42)
FX rate (% change)	−2276.6***	(−44.99)	−2083.1***	(−42.22)				
Unemployment	6.176***	(24.91)	5.488***	(24.11)				
Economic sentiment	−0.0256***	(−21.08)	−0.0056***	(−4.15)				
Balance of Payments (% GDP)	−42.082***	(−16.28)	−7.644***	(−28.39)				
FX rate volatility					122.6***	(12.11)	36.8**	(2.97)
10-year government bond yield					25.43***	(14.63)	2.31	(0.92)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	1,741,707		1,741,707		178,618		178,618	
Firms	429,978		429,978		48,470		48,470	
Distressed firms	30,900		30,900		4278		4278	
Log likelihood	−127,673.50		−59,226.49		−15,147.90		−5,004.10	
Wald test	55,783.8***		57,385.55***		8083.9***		9835.47***	

function. A hazard model is a type of survival model, in which the covariates are related to the time that passes before some event occurs (in this case distress). Specifically, we follow Laerkholm-Jensen et al. (2015) and estimate a fully parametric model with a constant baseline intensity, since the usual Cox semi-parametric model does not allow us to simultaneously identify the vector of macroeconomic coefficients as well as the time-varying baseline intensity.

Table 8 reports the estimation results for the overall sample and for the regional sub-samples under both techniques for comparison purposes. All our results remain robust when we instead apply the Cox model. Specifically, all regression coefficients retain their sign and significance. The sizes of the coefficients are very similar as well.

5.3. Time periods

We also split our sample into four rolling window periods (each one lasting five years). We 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.

In this section, we estimate the generic model for the overall sample 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. This time, we evaluate predictive power over exactly the next year following each model's

Table 9
Subperiods' Analysis (8 countries). The models are estimated over different sub-periods (five-year rolling windows for 2002–2009 data) with lagged 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). The data set includes non-financial SMEs from eight European economies. Standard errors are firm clustered-corrected (Huber/White standard errors). Parameter estimates are given first followed by chi-square values in parentheses. * denotes significance at a 5% level, ** at a 1% level and *** at a 0.1% level. FX rate is calculated in relation to the US dollar (USD/national currency). Unemployment is measured in%. The economic sentiment indicator is calculated by the Directorate General of Financial Affairs of the European Commission as an index with a mean value of 100, from answers to surveys conducted under the Joint Harmonized EU Program.

	2002–2006		2003–2007		2004–2008		2005–2009		2000–2009	
Panel A. Estimation results										
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)
FX rate (% change)	−1421.8***	(−29.32)	−1452.5***	(−33.02)	−478.9***	(−13.02)	−541.8***	(−14.16)	−1689.9***	(−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	Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes	
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***	
Panel B. Performance over next year										
Hosmer–Lemeshow decile										
1–5	8.05%		11.84%		11.94%		–		–	
8	14.43%		17.56%		13.18%		–		–	
9	19.67%		20.34%		18.99%		–		–	
10	44.06%		36.30%		40.75%		–		–	
8–10	78.15%		74.20%		72.93%		–		–	
Area under the ROC curve	0.818		0.783		0.796		–		–	
Panel C. Performance over last year (2009)										
Hosmer–Lemeshow decile										
1–5	12.80%		12.25%		11.94%		–		–	
8	16.34%		16.19%		13.18%		–		–	
9	18.45%		18.91%		18.99%		–		–	
10	35.86%		36.82%		40.75%		–		–	
8–10	70.65%		71.93%		72.93%		–		–	
Area under the ROC curve	0.780		0.783		0.796		–		–	

estimation period as well as over the last year of our sample (2009).

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 (period 2000–2009) 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 strongly from 2002 to 2003 (Table 2) but insolvency regimes remain stable or improve.

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

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, 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 the 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 in 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

three deciles ranges from 72.93% to 78.15% and AUC ranges from 0.7825 to 0.8177. Similarly, in Panel B, the percentage of distressed SMEs in the last three deciles ranges from 71.93% to 72.93% and AUC ranges from 0.7795 to 0.7963.

6. 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. We find that (in addition to financial indicators whose importance is noted in past studies) the location and number of shareholders are important determinants of SMEs' distress probabilities. 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. We find that as SMEs become larger, they are less vulnerable to the macroeconomic situation, contrary to what Basel regulations assume. Interestingly, when we split our sample in regional sub-samples, we show that SMEs across Europe are sensitive to the same firm-specific factors, but we identify significant regional variations in the selection and importance of macro variables. Specifically, macro variables differ among European regions based on region-specific conditions and characteristics. Since our regional distress models always perform better than a generic model estimated for the regional sub-samples, we conclude that their use can lead to performance improvements in the risk management of international SME portfolios. Finally, we perform a variety of tests and show that our

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 from an insolvent firm in the 2000–2009 period, the second column measures the average years from the insolvency filing until the resolution of assets and the third column is the ratio of the two previous columns. Data are from the World Bank and the 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

Appendix C

Distress statistics using main and alternative distress definitions. The appendix summarizes the properties of our three (one main and two alternative) distress indicators for the overall sample, the three regional sub-samples and the eight countries. It gives the number of total SMEs and distressed SMEs and the distress rate. Group 1 includes France, Germany and the UK, group 2 includes Italy, Portugal and Spain, and group 3 includes Czech Republic and Poland. According to the main distress definition, a firm-year is distressed if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) appears with one of the statuses "defaulted", "in receivership", "bankrupt", "in liquidation" or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. In the alternative distress definition 1, we exclude all firms that disappear from the sample before 2010 without updated status situation. These include firms that, under the main distress definition are classified as distressed if their equity is negative in the last year. Thus, the alternative distress definition 1 is strictly linked to a legal insolvency procedure. In the alternative distress definition 2, we exclude all firms that have negative equity in one or more of the years of their existence in the sample. Thus, under the alternative distress definition 2, we essentially lose some distress-related information because we include in the sample only firms with non-negative equity.

	Firm-years			Firms			Distressed			% of firm-years			% of firms		
	Main	Alter. 1	Alter. 2	Main	Alter. 1	Alter. 2	Main	Alter. 1	Alter. 2	Main	Alter. 1	Alter. 2	Main	Alter. 1	Alter. 2
Panel A. Overall sample and groups															
Overall sample	2,721,861	1,594,433	2,245,724	644,234	389,347	520,636	49,355	12,362	6836	1.81	0.78	0.30	7.66	3.18	1.31
Group 1	801,536	332,547	686,976	165,786	66,306	140,196	14,177	5646	3073	1.77	1.70	0.45	8.55	8.52	2.19
Group 2	1,741,707	1,185,258	1,412,358	429,978	302,959	341,471	30,900	6338	3566	1.77	0.53	0.25	7.19	2.09	1.04
Group 3	178,618	76,628	146,390	48,470	20,082	38,969	4278	378	197	2.40	0.49	0.13	8.83	1.88	0.51
Panel B. Countries															
Germany	21,681	5322	19,362	5954	1326	5271	319	8	6	1.47	0.15	0.03	5.36	0.60	0.11
France	724,060	309,230	624,182	145,918	61,030	124,300	12,222	5353	2867	1.69	1.73	0.46	8.38	8.77	2.31
United Kingdom	55,795	17,995	43,432	13,914	3950	10,625	1636	285	200	2.93	1.58	0.46	11.76	7.22	1.88
Italy	278,630	209,924	249,794	89,666	71,348	80,345	2257	219	83	0.81	0.10	0.03	2.52	0.31	0.10
Portugal	487,664	402,898	374,170	148,645	123,193	111,597	10,396	3702	1831	2.13	0.92	0.49	6.99	3.01	1.64
Spain	975,413	572,436	788,394	191,667	108,418	149,529	18,247	2417	1652	1.87	0.42	0.21	9.52	2.23	1.10
Czech Republic	119,677	59,856	93,746	33,305	16,804	25,631	3014	244	111	2.52	0.41	0.12	9.05	1.45	0.43
Poland	58,941	16,772	52,644	15,165	3278	13,338	1264	134	86	2.14	0.80	0.16	8.33	4.09	0.64

results remain robust to different distress definitions, estimation techniques and time periods.

Appendix A

See Appendix Tables A–C.

References

Acharya, Viral V., Sundaram, Rangarajan K., John, Kose, 2011. Cross-country variations in capital structure: the role of bankruptcy costs. Journal of Financial Intermediation 20 (1), 25–54.
Allen, Linda, DeLong, Gayle, Saunders, Anthony, 2004. Issues in the credit risk modeling of retail markets. Journal of Banking and Finance 28 (4), 727–752.
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., Iwanicz-Drozdowska, Malorzata, Laitinen, Erkki K., Suvas, Arto, 2014. Distressed firm and bankruptcy prediction in an international context: a review and empirical analysis of Altman's Z-score model. Working Paper.
Altman, Edward I., Hotchkiss, Edith, 2005. Corporate Financial Distress and Bankruptcy, third ed. Wiley.
Altman, Edward I., Rijken, Herbert, 2011. Toward a bottom-up approach to assessing sovereign default risk: an update. Journal of Applied Finance 21 (2), 5–19.
Altman, Edward I., Sabato, Gabriele, 2007. Modelling credit risk for SMEs: evidence from the U.S. market. Abacus 43 (3), 332–357.
Altman, Edward I., Sabato, Gabriele, Wilson, Nicholas, 2010. The value of non-financial information in SME risk management. Journal of Credit Risk 6 (2), 95–127.
Becchetti, Leonardo, Sierra, Jaime, 2003. Bankruptcy risk and productive efficiency in manufacturing firms. Journal of Banking and Finance 27 (11), 2099–2120.
Beck, Thorsten, Demirguc-Kunt, Asli, Maksimovic, Vojislav, 2008. Financing patterns around the world: are small firms different? Journal of Financial Economics 89 (3), 467–487.
Berger, Allen N., Udell, Gregory F., 2006. A more complete conceptual framework about SME finance. Journal of Banking and Finance 30 (11), 2945–2966.
Bonfim, Diana, 2009. Credit risk drivers: evaluating the contribution of firm level information and of macroeconomic dynamics. Journal of Banking and Finance 33 (2), 281–299.
Bruneau, Catherine, de Bandt, Olivier, El Amri, Widad, 2012. Macroeconomic fluctuations and corporate financial fragility. Journal of Financial Stability 8 (4), 219–235.
Campbell, John Y., Hilscher, Jens, Szilagyi, Jan, 2008. In search of distress risk. The Journal of Finance 63 (6), 2899–2939.
Carling, Kenneth, Jacobson, Tor, Lindé, Jesper, Roszbach, Kasper, 2007. Corporate credit risk modeling and the macroeconomy. Journal of Banking and Finance 31 (3), 845–868.
Chava, Sudheer, Jarrow, Robert, 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.
Creditreform, 2012 Insolvencies in Europe. Surveys by the Creditreform Economic Research Unit, available at <http://www.creditreform.com/en/press-and-downloads/analysis.html>.
Davydenko, Sergei A., 2012. When do firms default? A study of the default boundary. Working Paper.
Davydenko, Sergei A., Franks, Julian R., 2008. Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the U.K. The Journal of Finance 63 (2), 568–608.
Dietsch, Michel, Petey, Joël, 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, Saita, Leandro, Wang, Ke, 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, Hayden, Evelyn, Tasche, Dirk, 2003. Testing rating accuracy. Journal of Risk 16 (1), 82–86.
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., Vetsuypens, Michael R., 1993. CEO compensation in financially distressed firms: an empirical analysis. The Journal of Finance 48 (2), 425–458.
Glennon, Dennis, Nigro, Peter, 2005. Measuring the default risk of small business loans: a survival analysis approach. Journal of Money, Credit and Banking 37 (5), 923–947.
Gormley, Torr A., Matsa, David A., 2014. Common errors: how to (and not to) control for unobserved heterogeneity. Review of Financial Studies 27 (2), 617–661.
Grammatikos, Theoharry, Vermeulen, Robert, 2012. Transmission of the financial and sovereign debt crises to the EMU: stock prices, CDS spreads and exchange rates. Journal of International Money and Finance 31 (3), 517–533.
Grunert, Jens, Norden, Lars, Weber, Martin, 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.

Hosmer, David W., Lemeshow, Stanley, 2000. Applied Logistic Regression, second ed. Wiley.
Jacobson, Tor, Lindé, Jesper, Roszbach, Kasper, 2005. Exploring interactions between real activity and the financial stance. Journal of Financial Stability 1 (3), 308–341.
Jacobson, Tor, Lindé, Jesper, Roszbach, Kasper, 2013. Firm default and aggregate fluctuations. Journal of the European Economic Association 11 (4), 942–972.
Jaumotte, Florence, Sodsriwiboon, Piyaporn, 2010. Current account imbalances in the southern euro area. IMF Working Paper WP/10/139.
Laerkholm-Jensen, Thais, Lando, David, Medhat, Mamdouh, 2015. Cyclicality and firm size in private firm defaults. Working Paper.
LaPorta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, Vishny, Robert W., 1998. Law and Finance. Journal of Political Economy 106 (6), 1113–1155.
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, Kim, Tong Suk, Park, Nam Jung, Lee, Hoe Kyung, 2008. Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. Journal of Forecasting 27 (6), 493–506.
Perego, Erika R., Vermeulen, Wessel N., 2013. Macroeconomic determinants of European stock and government bond correlations: a tale of two regions. Working Paper.
Ross, Stephen A., Westerfield, Randolph, Jaffe, Jeffrey, 2010. Corporate Finance, ninth 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, Everett, Jim, 1993. Defining small business failure. International Small Business Journal 3 (11), 35–48.
White, Halbert, 1994. Estimation, Inference and Specification Analysis. Cambridge University Press.
Wooldridge, Jeffrey M., 2002. Econometric Analysis of Cross-section and Panel Data. MIT Press.