DANMARKS NATIONALBANK

DANMARKS NATIONALBANK WORKING PAPERS 2005 • 26

Anne Dyrberg Rommer

Financial Markets
Danmarks Nationalbank, Copenhagen
and
Centre for Applied Microeconometrics (CAM),
Institute of Economics, University of Copenhagen

A Comparative Analysis of the Determinants of Financial Distress in French, Italian and Spanish firms

18 May 2005

Please direct any enquiries to
Danmarks Nationalbank, Information Desk, Havnegade 5, DK-1093 Copenhagen K
Denmark
Tel.: +45 33 63 70 00 (direct) or +45 33 63 63 63
Fax : +45 33 63 71 03
E-mail:info@nationalbanken.dk

Henvendelser kan rettes til :
Danmarks Nationalbank, Informationssektionen, Havnegade 5, 1093 København K.
Telefon: 33 63 70 00 (direkte) eller 33 63 63 63
E-mail: info@nationalbanken.dk

ISSN (trykt/print) 1602-1185

ISSN (online) 1602-1193

Resumé

De faktorer, der fører til økonomiske vanskeligheder i italienske, spanske og franske små og mellemstore virksomheder, analyseres i et forsøg på at finde ud af, om det er de samme eller om det er forskellige faktorer, der fører til økonomiske vanskeligheder i landene. For at kunne foretage analysen estimeres individuelle kreditrisikomodeller for alle tre lande. Til analysen benyttes et datasæt, som er leveret af Bureau van Dijk. Det fantastiske ved datasættet er, at det muliggør sammenligning på tværs af landene. En minus ved datasættet er imidlertid, at det ikke er lige så godt, som nogle af de datasæt, der er brugt i individuelle lande studier (forstået på den måde, at et antal virksomheder dropper ud af panelet uden forklaring).

Sammenligningen af signifikans-niveauet og fortegnet af parameter estimaterne viser, at på trods af, at der er visse ligheder mellem landene, så er der også store forskelle mellem landene. Afkastningsgraden og solvensgraden er nogle af de variabler, der påvirker landene på samme måde. De variabler, der ikke er signifikante i alle lande eller har forskelligt fortegn, er lån i forholdet til aktiverne, størrelse, alder, juridisk form og stor koncentration af ejerskab.

Foruden de individuelle kreditrisikomodeller estimeres en samlet kreditrisikomodel for alle tre lande. Da estimater af sandsynligheden for at ende i økonomiske vanskeligheder kræver en stor mængde data, tillader Basel II, at banker slår deres data sammen med andre banker. Herved er et antal af internationale data projekter opstået, hvor banker fra forskellige lande slår deres data sammen. På grund af denne udvikling, og da, endvidere, mange kreditinstitutioner i Europa har engagementer i flere lande, da er valget mellem en individuel kreditrisikomodel og en fælles kreditrisikomodel relevant, når kapitalkravene skal regnes ud for bankerne. Sammenligningen af de signifikante variabler og deres fortegn og forudsigelsesevnen i de tre landemodeller og den fælles kreditrisikomodel viser, at resultaterne i den fælles model er ret forskellige fra resultaterne i landemodellerne.

Der er få studier, der på tværs af lande sammenligner de faktorer, der fører til økonomiske vanskeligheder. Så vidt vi ved, er dette det første komparative studie af kreditrisikomodeller estimeret for en forholdsvis homogen gruppe af lande. Dermed udfylder studiet et hul i litteraturen.

Abstract

The determinants of corporate failure in Italian, Spanish and French small and medium-sized enterprises are investigated in order to find out whether the predictors of financial distress in the countries are the same or not. In order to compare the determinants of financial distress, accounting-based credit-scoring models for each country are estimated. The analysis uses a data set provided by Bureau van Dijk. The great virtue of the data set is that it enables us to make cross-country comparisons. On the negative side it should be mentioned that the data set, when looking at each country individually, is not as good as some of the data sets used in the individual country studies (in the sense that a number of firms drop out of the panel with no explanation).

The comparison of the significance and sign of the determinants of financial distress in the three countries shows, that although there are some similarities across countries, there are also quite a lot of differences. Some of the variables that behave similarly across countries are the earnings ratio and the solvency ratio. The variables, whose effect differs between the countries in terms of whether or not they are significant or what sign they have, are the loans to total assets ratio, size, age, legal form and a variable, which measures the concentration of ownership.

Apart from the individual credit-scoring models a model including all countries is estimated. As valid estimates of the probability of default for individual banks require a considerable amount of data, Basel II allows for banks to pool their data with other banks in order to overcome their data shortcomings. In this way a number of international data pooling projects have emerged, where banks from various countries pool their data. Because of this development and as, furthermore, many credit institutions in Europe have cross-border activities, the choice between setting up individual country credit-scoring models or a common credit-scoring model is relevant, when calculating capital requirements for banks. The comparison of the significance and sign of the parameter estimates and the predictive ability of the individual country credit-scoring models and the pooled model show that the pooled model delivers results that differ to quite an extent from the individual country credit-scoring models.

There are few studies, which do compare the determinants of financial distress across countries. To the best of our knowledge, this is the first comparative accounting-based credit-rating study of a fairly homogenous group of countries, and so it fills a gap in the literature.

ANNE DYRBERG ROMMER*

A COMPARATIVE ANALYSIS OF THE DETERMINANTS OF FINANCIAL DISTRESS IN FRENCH, ITALIAN AND SPANISH FIRMS

This version: 18 May 2005

* The author would like to thank Hans Christian Kongsted, Ivan Alves, Jesper Berg, Karsten Biltoft, Olli Castren, Anders Møller Christensen, Frank Dierick, John Fell, Kenneth Juhl Pedersen and Bronka Rzepkowski for commenting on earlier versions of the paper and participants at the European Central Bank seminar, which took place on the 23rd of February 2005. Corresponding address is: Anne Dyrberg Rommer, Financial Markets, Danmarks Nationalbank, Havnegade 5, DK-1093 Copenhagen, Denmark. Phone: + 45 33 63 63 63. Email: ady@nationalbanken.dk.

Table of contents

A COMPARATIVE ANALYSIS OF THE DETERMINANTS OF FINANCIAL DISTRESS IN FRENCH, ITALIAN AND SPANISH FIRMS

1. Introduction

The determinants of corporate failure in Italian, Spanish and French small and medium-sized enterprises (SMEs) are investigated in order to find out whether the predictors of financial distress in the countries are the same or not. In order to compare the determinants of financial distress, accounting-based credit-scoring models for each country are estimated. An accounting-based credit-scoring model is a model, which on the basis of information extracted from company accounts, and perhaps also non-financial information (such as the age of the company), estimates the probability that a particular firm will default on its debt obligations, usually over a one year horizon. Furthermore a model including all countries is estimated. The significant variables and their sign, and its predictive ability is compared to the three country models in order to assess the differences in the determinants of financial distress and in the predictive ability of the two model set ups. There are few studies, which do compare the determinants of financial distress across countries. To our knowledge, this is the first comparative accounting-based credit-rating study of a fairly homogenous group of countries, and so it fills a gap in the literature.

Some of the few studies, which do compare the determinants of failure in several countries, are Hunter and Isachenkova (2000), Bhattacharjee, Higson, Holly and Kattuman (2004) and Ooghe and Balcaen (2002). Hunter and Isachenkova (2000) aim at explaining the differences in predictors of failure in Russian and UK industrial firms. They pick Russian and UK industrial firms, as the results may be useful for the increasing number of western businesses, which export to Russia etc., or for governmental bodies in Russia and international agencies, which provide assistance to Russia. They find that liquidity and gearing are not effective in explaining failure in Russian companies, whereas measures such as size, profitability and turnover seem to be robust predictors. The UK results indicate the importance of profitability, gearing and liquidity. Bhattacharjee et al. (2004) analyse UK and US quoted firms. They set up a competing-risks model to identify the characteristics leading to bankruptcy and acquisition, and they find that there are significant differences in the way in which firms in the UK and US react to changes in the macroeconomic environment. They argue that these differences in response may be attributable to differences in bankruptcy codes in the UK and the US. Another comparative study is Ooghe and Balcaen (2002). Their focus is on whether a given failure prediction model can be easily transferred across countries. Failure prediction models from different countries are

compared using a dataset of Belgian company accounts. Their study can be seen as a case study on the "transferability" of models developed in a specific country and period to other countries and/or periods.

Along the lines of Hunter and Isachenkova (2000) and Bhattacharjee et al. (2004) the purpose here is to discuss possible differences and similarities in the determinants of failure in firms in different countries. In contrast to these two other studies, this paper focuses on countries that in important aspects are fairly alike. They all belong to Continental Europe, are a part of the European Monetary Union and they are inspired by the same legal tradition. Furthermore, despite the deregulation and liberalisation process of the financial systems, which took place in the countries in the 1980s and 1990s, banks are still very important sources of financing in all three countries.[1] As the countries are fairly alike, a priori, one may think that the same factors drive financial distress in the three countries. One could even argue that if common factors should drive financial distress across countries, it would be in countries that are alike in so many important aspects, such as Spain, Italy and France.

The analysis has implications for at least two policy areas. The first concerns financial stability analysis and the second the Revised Framework for Capital Measurement and Capital Standards, also known as Basel II. An important part of financial stability analysis entails assessing the degree of corporate sector credit risk facing banks.[2] For financial stability analysis on a euro area wide basis, it is important to ascertain whether common or country-specific factors drive corporate failures. If the factors that give rise to financial distress are the same across countries, then aggregation of individual corporate sectors into a single group is justified, whereas, if country specific factors are more important, this would call for analysing conditions in each individual corporate sector. Basel II opens up for the possibility that credit institutions themselves can estimate their minimal capital requirements. According to Basel II credit institutions can choose between one of two internal ratings based approaches when they calculate their capital requirements.[3] If they choose to do so and follow one of the two internal ratings based approaches they have to calculate the probability of default of their obligors in order to calculate their minimal capital requirements.[4] As valid estimates of the probability of default require a considerable amount of data, Basel II allows for banks to pool their data with other banks in order to overcome their data shortcomings, and so a number of international data pooling projects have emerged, where banks from various countries pool their data, c.f. Borup, Kurek and Rommer (2005). Because of this development and as, furthermore, many credit institutions in Europe have cross-border activities, the choice between setting up individual country credit-scoring models or a common credit-scoring model

[1] See ECB (2002).

[2] Financial stability analysis of non-financial firms usually involves examining conditions in small and medium-sized enterprises (SMEs) as well as large companies separately. In order to assess the financial health of large companies, a number of information sources are available, such as credit-ratings and market-based indicators such as expected default frequencies. These sources are not available for most SMEs. Instead, the analysis of SMEs usually relies on company accounts. For this study, income statement and balance sheet information was collected for SMEs in France, Italy and Spain.

[3] For further details, see Basel Committee on Banking Supervision (2004).

[4] This study follows the European Commission definition of SMEs, which differs from the definition of SMEs in Basel II.

is relevant, when calculating capital requirements for banks. In order to shed light on these issues, the determinants of corporate failure in French, Italian and Spanish firms are investigated by the estimation of individual country credit-scoring models and a common credit-scoring model.

The paper is structured as follows. Section 2 presents the literature. A discussion of data is found in section 3, which also sets up the hypotheses to be tested. Section 4 presents the econometric theory, and the results are discussed in section 5. Section 6 compares the individual country models with the pooled model. Section 7 concludes.

2. The Literature

Credit-scoring studies using French, Italian and Spanish data are listed in section 9 (appendix). The studies using French data are the following: Bardos (1998) and Bardos (2001) present the Banque de France's scoring method and discuss the recent developments of the method. Dietsch and Petey (2002) propose a credit model for French SME loans. Moody's Investors Service (2001b) sets up a country model using data from France. A number of studies use Spanish data. The focus in Jiménez and Saurina (2004) is on the impact of certain loan characteristics on credit risk, e.g. collateral, type of lender institutions and the relationship between the bank and the company it is financing. Corcóstegui, González-Mosquera, Marcelo and Trucharte (2003) estimate a rating system for the Spanish non-financial private-sector firms. Fernandez (1988) estimates a Spanish model for credit risk classification. Moody's Investors Service (2001a) sets up a country model using data from Spain. There are several studies using Italian data. Cifarelli and Corielle (1988) describe an application of a Baysian variant to discriminant analysis. Altman, Marco and Varetto (1994) analyze the comparison between traditional statistical methodologies for distress classification and prediction with an artificial intelligence algorithm (neural networks). Moody's Investors Service (2002) sets up a country model using data from Italy.

All the studies differ in terms of sample selection procedure (including period covered and default definition) and econometric technique. None of them are of a comparative nature. In fact, very few studies compare the determinants of failure across countries. As mentioned in the introduction, some of the few studies, which do compare the determinants of failure in several countries, are Hunter and Isachenkova (2000) and Bhattacharjee et al. (2004). The studies, which analyse Russian firms and UK firms and firms in the UK and US, respectively, differ from this study, whose focus is on a fairly homogenous group of countries.

Various estimation techniques have been suggested in the accounting-based credit-scoring literature (e.g. discriminant analysis and the logit model).[5] When estimating a credit-scoring model, typically, information on two groups of firms is gathered and used in the estimations, namely information on

[5] For an overview of the literature the reader is referred to Lando (2004). Some of the often quoted accounting-based credit-scoring studies are Beaver (1966), Altman (1968), Ohlson (1980) and Shumway (2001).

firms in financial distress and active firms. Along the lines of Dyrberg (2004), this paper extends the common practice in credit-rating studies and it includes also firms that exit for other reasons than financial distress (e.g. voluntary liquidations). As firms can exit for various reasons a competing-risks model is set up. The estimation strategy of Allison (1982) is followed and the probability of exiting to the various states is estimated simultaneously. Methodologically related papers are Harhoff, Stahl and Woyde (1998), Köke (2001), Prantl (2003) and Bhattacharjee et al. (2004). All four studies distinguish between two forms of exit. Harhoff et al. (1998) and Prantl (2003) model voluntary liquidations and bankruptcies in Germany using a competing-risks framework. Bhattacharjee et al. (2004) use a competing-risks model to identify the characteristics leading to bankruptcy and acquisition in UK and US quoted firms. Köke (2001) investigates the determinants of acquisition and failure in Germany. He provides stylized facts and discusses lessons for empirical studies of firms.

3. Data and the Hypotheses to be tested

This section discusses the dataset used in the estimations, including the construction of the dependent variable and the sample selection criteria, and it sets up hypotheses to be tested.

3.1 Construction of the Dependent Variable

The data used for Italy, Spain and France comes from the harmonised Amadeus database, which is a pan-European database, provided by Bureau van Dijk (BvD).[6] As opposed to most Italian, Spanish and French credit-scoring models presented in the literature, which use non-public information from credit registries operated by governments (usually by bank supervisors) or from other non-public sources, c.f. section 9 (appendix), this study uses public information only. The great virtue of the data set is that it enables us to make cross-country comparisons. On the negative side it should be mentioned that the data set, when looking at each country individually, is not as good as some of the data sets used in the individual country studies (in the sense that a number of firms drop out of the panel with no explanation).

Amadeus contains information on public and private companies. A standard company report includes balance sheet items, profit and loss account items and non-financial information such as BvD ID number, address, legal status, date of incorporation, sector affiliation code, number of employees,

[6] The data is loaded from the Amadeus cd-roms dated September 2004, October 2004 and November 2004. Working with the Amadeus database has delivered one important by-product for other studies taking their point of departure in the database. We investigated whether or not the analysis could be done for Germany, but it could not. The reason is that Creditreform, who provides German data to Amadeus, only delivers company accounts for the "best" German firms. Creditreform divides the German financial accounts into 6 different risk classes. The four best classes are included in Amadeus. The two worst classes are not included. These last two classes include defaulting firms, and so, if one wants to estimate a credit-scoring model, one cannot do this from the data in Amadeus only. This finding, which is not described in the Amadeus documentation, is important for other studies as well. One needs to be aware of the fact that only relatively "good" German financial accounts can be obtained from Amadeus. The sample is not representative.

ownership information etc. Issues regarding data and the sample selection procedures are discussed in the following sections, and the hypotheses to be tested are set up.

The focus in this paper is on the firms that end up in financial distress, or in other words, firms that can be expected to inflict a loss on the financial sector. In the measure of financially distressed firms the following events are included: 1) bankruptcy, 2) active (receivership) and 3) active (default of payments).[7] The firms in financial distress are referred to as E1 exits. Dissolved (merger) are denoted E2 exits and voluntary liquidations are denoted E3 exits. Inactive (no precision) are denoted E4. These firms are known to exit the database, but it is unknown what the reason for the exit is. Our best presumption is that these firms exit for other reasons than E1, E2 and E3 reasons. It is sure that they do not exit as E5 firms, as E5 firms are active firms that hand in a financial statement in 2000 and 2001, but not in 2002, or in 2000 only. The E5 firms constitute a residual category. It is not known whether these firms end up as E1, E2, E3, E4 or active firms. The E5 firms are called attritioners. They are discussed in great length in section 3.3.2. The reference group is active firms. By construction, all exits are equal to E1+E2+E3+E4+E5. Table 3.1.a provides an overview of the various exits.

Table 3.1.a: The construction of the dependent variable

Code in database	Type of exit	Category	France	Italy	Spain
Active	Reference group		X	X	X
Bankruptcy	Financial distress	E1	X	X	X
In liquidation	Voluntary exit	E3	X	X	Not reported
Dissolved (merger)	Merger	E2	X	X	X
Active (receivership)	Financial distress	E1	X	Not reported	Not reported
Inactive (no precision)	Unknown why the firms have exited	E4	X	X	X
Active (default of payments)	Financial distress	E1	X	Not reported	X
Dissolved	Voluntary exit	E3	Not reported	X	X
Self-constructed category: Active firms that hand in a financial statement in 2000 and 2001, but not in 2002. Active firms that hand in a financial statement in 2000, but not in 2001 and 2002	Unknown whether the firms has exited or if it is still active	E5	X	X	X

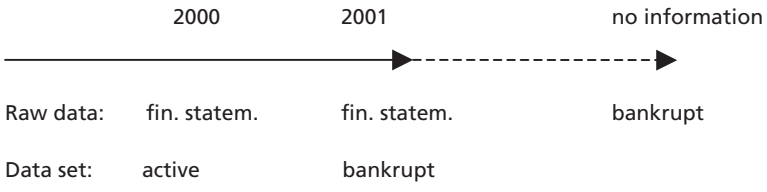
Note: See tables 10.a, 10.b and 10.c in section 10 (appendix) for country specific details of the legal status codes.

[7] The definition of firms in financial distress differs from the definition used in the Revised Framework for Capital Measurement and Capital Standards, also known as Basel II, where the obligors 1) that are past due more than 90 days on any credit obligation or 2) those that, with a high probability, can be considered unable to pay their credit obligations, are defined as firms that are defaulting (at least one of the conditions must be met), c.f. Basel Committee on Banking Supervision (2004). The event "financially distressed" is a fairly late credit event compared to the Basel II definition. It was not possible to follow the Basel II definition, as this model is based on public information only (and no bank default data is available). This does not seem to make a big difference, when building the credit-scoring model. Hayden (2003) shows that credit-scoring models that rely on bankruptcy as default criterion instead of delay-in-payments can be equally powerful in predicting the credit loss events. Furthermore, Moody's Investors Service (2001a) reports that experience shows that the factors that can predict default are generally the same, no matter whether the definition of defaults is 90 days past due or bankruptcy.

Unfortunately, information on the firms that exit (or, in technical terms, the legal status variable) is only kept in the database for 3 years, and so, currently, the estimations of the credit-scoring models cover only firms that have handed in financial statements in the period 2000 – 2002.[8] Firms that hand in a financial statement in 2000 are recorded as belonging to year 2000. Firms that hand in a financial statement in 2001 are recorded as belonging to year 2001 etc.

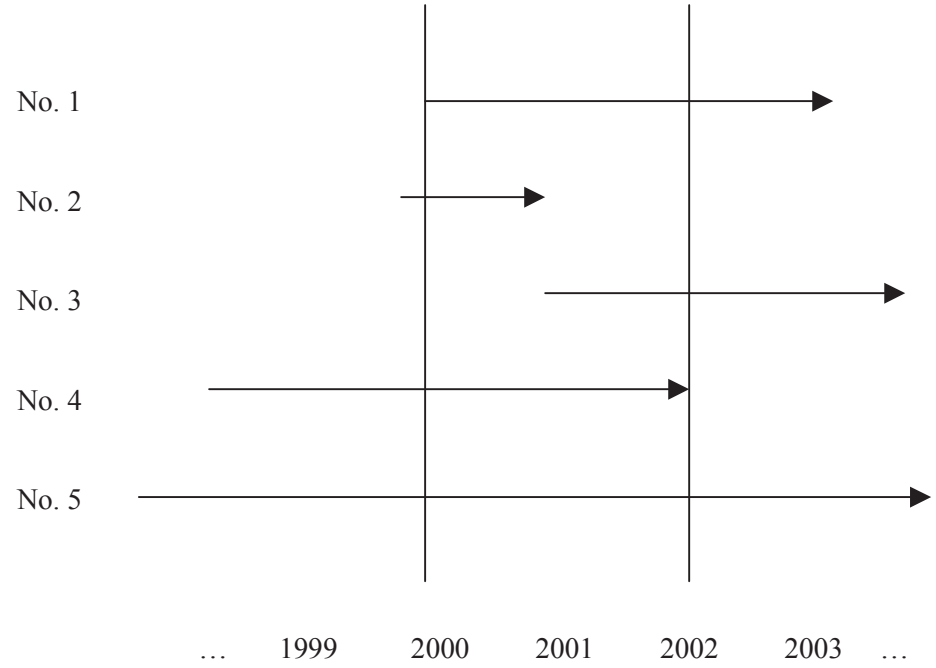
There is a lag between a firm's last financial statement and the registration of the legal status. The timing of events could follow the time line sketched in figure 3.1.b. In the figure, the firm is active in 2000 and 2001, and it hands in a financial statement both years. Later it goes bankrupt, and therefore the firm is registered with a bankruptcy code, c.f. the line sketching the "raw data" in figure 3.1.b. To use this information in the estimations, a recoding of the dataset has taken place, c.f. the line "dataset" in figure 3.1.b: In the dataset the firm is coded to be active in 2000 and bankrupt in 2001. The firm in figure 3.1.b corresponds to the firm, which is represented by spell no. 2 in figure 3.1.c. The interpretation of the construction of the data is that the "actual" time of the bankruptcy is the day after the firm hands in its last financial statement. In reality the "actual bankruptcy" happens at some unknown point in time, i.e. it is not registered by Bureau van Dijk. For the analysis the exact bankruptcy date is not important. The amount of time that passes between the time the specific firm hands in its last financial statement in 2001 until it goes bankrupt is arbitrary anyway, as it depends on the bankruptcy court that handles the specific case (how many cases it has already, etc).

Figure 3.1.b: A firm's last financial statement and the registration of the codes



[8] Ideally the estimation period would cover a whole business cycle. A future extension of this work could be to try and collect more data, or at least to save the data on financial distress as time moves along, as another year and yet another year, can then be incorporated into the dataset.

Figure 3.1.c: Observation window (5 examples of firms that are used in the analysis)



Note: The data set consists of a single spell for each firm, meaning that once a firm has exited it cannot re-enter. This figure pictures five different spells (see the arrows in the figure). Each firm is observed when it hands in its financial statement in 2000, 2001 and 2002 (if it does not exit during the period). The period 2000 to 2002 is called the observation window. As the firms are observed once a year, but interact on a continuous basis, the data is called grouped duration data. The data set consists of firms that are both flow and stock sampled. Spell no. 1 and 3 are flow sampled and spell no. 2, 4 and 5 are stock sampled. The firms that enter the active state in the period of interest (in 2000, 2001 or 2002) are flow sampled. The firms that are already active in 2000 are stock sampled. The firms are censored in various ways (as they are only observed in the observation window). Stock sampled firms can be left censored, left truncated and right censored. They are left censored, if the incorporation data is unknown. They are left truncated, if the incorporation data is known. The firms that survive beyond the observation window are right censored, as they are not observed after the end of the observation window. The flow sampled firms are not left censored (or left truncated) as their history is observed from the beginning of the spell. The flow sampled firms that survive beyond the observations window are right censored, as they are not observed after the end of the observation window.

3.2 Hypotheses

A number of hypotheses, that are to be tested, are set up in this section. The focus is on the way various factors may affect the likelihood of exiting as an E1 firm. No hypothesis will be set up on the effect of the institutional framework. Spain, Italy and France are all French-civil law countries[9], and so large differences between the institutional frameworks are not present, c.f. table 3.2.a., which shows how the countries do on enforcement variables (e.g. “efficiency of judicial system”, “rule of law” and

[9] La Porta et al. (1998) discuss the rules and practices governing the resolution of financial distress in 49 countries. Their focus is on company and bankruptcy/reorganization laws. They explain how commercial laws come from two broad traditions. One tradition is the common-law family, which is English in origin. The other tradition is civil law, which derives from Roman law. Within the civil tradition, the modern commercial laws can have French, German, and Scandinavian origin. The French Commercial Code, which Italy, Spain and France are inspired by, dates back to Napoleon in 1807. The German Commercial code was written in 1897 after Bismarck’s unification of Germany. The Scandinavian laws are “similar to each other but “distinct” from others” (La Porta et al. (1998:1119)). This study focuses on France, Italy and Spain, which are inspired by the French Commercial Code. Examples of credit-scoring studies that take their point of departure in the other law families are Bunn and Redwood (2003) (common-law family), Scheule (2003) (German origin) and Dyrberg (2004) (Scandinavian origin).

"corruption"), accounting standards and creditor rights (e.g. "no automatic stay on assets", "secured creditors paid first"). Table 3.2.a shows, that the three countries do better on the enforcement variables and on the rating on accounting standards than the French-civil average (which is calculated on the basis of all French-civil law countries studied in La Porta et al. (1998) and not only on the basis of France, Italy and Spain)[10], but, that France does worse than the French-civil average when measured on creditor rights[11]. These issues will not be addressed further. Instead it will be discussed how other factors, which possibly are affected by the rules and practices governing the resolution of customers' financial distress, may explain, why different determinants of financial distress in the countries end up being significant in the estimations. First, the common indicators in credit-scoring models – profitability, solvency, leverage, age and size – are discussed. Thereafter, various proxies, which are used for the inherently unobservable variables, are discussed. All variables are listed in table 3.2.b.

Hypothesis 1 concerns the effect of profitability (measured as the earnings ratio, here defined as EBITDA (earnings before interest, taxes, depreciation and amortization) to total assets). The hypothesis is that the sign of the coefficient to the earnings ratio is negative in all countries. *Hypothesis 2* concerns the solvency ratio, which expresses the firm's ability to generate satisfactory earnings over time. It is calculated as shareholders funds over total assets, and it can be seen as a buffer. The hypothesis for all countries is that a high solvency ratio decreases the probability of moving into financial distress. *Hypothesis 3* concerns the leverage measured as loans over total assets.[12] The hypothesis is that higher leverage increases the probability of moving into financial distress.

Hypothesis 4 and 5 concerns the size and the age of the firms. As these variables are correlated (e.g. because older firms tend to be larger than younger firms) it can be difficult to disentangle the effects stemming from age and size. *Hypothesis 4* concerns the size of a firm, which is measured as the logarithm of total assets. The hypothesis is that the larger the firm is, the less likely it is to enter financial distress. *Hypothesis 5*: Age is also included in the estimations. Dummies for each age are included (1 year old firms are chosen as the reference group). In the final results, only significant age dummies are presented. The hypothesis is that the older the firm is, the less likely it is to enter financial distress.

Hypothesis 6 concerns a proxy for diversification, namely a variable, which measures the number of subsidiaries attached to the specific company. It is left to the estimations to show whether or not there is a significant effect of this variable.

---

[10] Except one variable for Spain: The enforcement variable "efficiency of judicial system" is lower for Spain than the French-origin average.

[11] Note also that the legal reserve required as a percentage of capital in the three countries is smaller than the French-origin average.

[12] In the Amadeus database, current liabilities are split up on loans, creditors and other current liabilities. To construct the leverage measure, the loans item is used.

Table 3.2.a: Enforcement variables, rating on accounting standards and creditor rights (higher score is better enforcement, accounting standards and creditor rights, respectively)

		France	Italy	Spain	French-origin average
Enforcement variables	Efficiency of Judicial System	8.00	6.75	6.25	6.56
	Rule of Law	8.98	8.33	7.80	6.05
	Corruption	9.05	6.13	7.38	5.84
	Risk of Expropriation	9.65	9.35	9.52	7.46
	Risk of contract Repudiation	9.19	9.17	8.40	6.84
Accounting standards	Rating on accounting standards	69	62	64	51.17
Creditor rights (1=creditor protection is the law)	No automatic stay on assets	0	0	1	0.26
	Secured creditors first paid	0	1	1	0.65
	Restrictions for going into reorganization	0	1	0	0.42
	Management does not stay in reorganization	0	0	0	0.26
	Creditor rights, sum	0	2	2	1.58
	Legal reserve required as a percentage of capital	0.10	0.20	0.20	0.21

Note: The French-civil average consist of Argentina, Belgium, Brazil, Chile, Colombia, Ecuador, Egypt, France, Greece, Indonesia, Italy, Jordan, Mexico, Netherlands, Peru, Philippines, Portugal, Spain, Turkey, Uruguay and Venezuela. *Enforcement variables*: Efficiency of judicial system. Assessment of the "efficiency and integrity of the legal environment as it affects business, particularly foreign firms'" produced by the country risk rating agency Business International Corp. It "may be taken to represent investors' assessments of conditions in the country in question." Scale from zero to 10; with lower scores, lower efficiency levels. Rule of law. Assessment of the law and order tradition in the country produced by the country risk rating agency International Country Risk (ICR). Scale from zero to 10, with lower scores for less tradition for law and order. Corruption. ICR's assessment of the corruption in government. Lower scores indicate that "high government officials are likely to demand special payments" and "illegal payments are generally expected throughout lower levels of government" in the form of "bribes connected with import and export licences, exchange controls, tax assessment, policy protection, or loans." Scale from zero to 10, with lower scores for higher levels of corruption. Risk of expropriation. ICR's assessment of the risk of "outright confiscation" or "forced nationalisation". Scale from zero to 10, with lower scores for higher risks. Risk of contract repudiation. ICR's assessment of the "risk of a modification in a contract taking the form of a repudiation, postponement, or scaling down" due to "budget cut-backs, indigenization pressure, a change in government, or a change in economic and social priorities". Scale from zero to 10, with lower scores for higher risks. *Accounting standards.* Index created by examining and rating companies' annual reports on their inclusion or omission of 90 items. These items fall into seven categories (general information, income statements, balance sheets, funds flow statement, accounting standards, stock data, and special items). A minimum of three companies in each country were studied. *Creditor rights.* Restrictions for going into reorganisation equals one if the reorganisation procedure imposes restrictions, such as creditors' consent, to file for reorganisation; equals zero if there are no such restrictions. No automatic stay on secured assets equals one if the reorganisation procedure does not impose an automatic stay on the assets of the firm on filing the reorganisation petition. Secured creditors' paid first equals one if secured creditors are ranked first in the distribution of the proceeds that result from the disposition of the assets of a bankrupt firm; equals zero if non-secured creditors, such as the government and workers, are given absolute priority. Management does not stay equals one when an official appointed by the court, or by the creditors, is responsible for the operation of the business during reorganization. Equivalently, this variable equals one if the debtor does not keep the administration of its property pending the resolution of the reorganization process. Equals zero otherwise. Creditor rights is an index aggregating different creditor rights. The index is formed by adding 1 when (1) the country imposes restrictions, such as creditors' consent or minimum dividends to file for reorganisation; (2) secured creditors are able to gain possession of their security once the reorganization petition has been approved (no automatic stay); (3) secured creditors are ranked first in the distribution of the proceeds that results from the disposition of the assets of a bankrupt firm; and (4) the debtor does not retain the administration of its property pending the resolution of the reorganization. The index ranges from zero to four. Legal reserve. The minimum percentage of total share capital mandated by corporate law to avoid the dissolution of an existing firm. It takes a value of zero for countries without such a restriction. Source: La Porta et al. (1998)

Table 3.2.b: The explanatory variables used in the study

	Variable
Core variables	Profitability: Earnings ratio= EBITDA/total assets. EBITDA = earnings before interest, taxes, depreciation and amortization
	Solvency: Solvency = Shareholders funds/total assets
	Leverage: Loans/total assets
	Firms size: Log(total assets)
	Age: Dummies. Reference category is firms that are 1 year old.
Proxies	Subsidiaries: This variable measures the number of subsidiaries that a company has registered.
	Legal form: This dummy is equal to 1, if it is a private limited liability company, and equal to 0, if it is a public limited liability company.
	Shareholders: This variable measures the number of recorded shareholders.
	Independence indicator: Three dummies are included. The first is equal to 1 when it is a type A firm (bvd_indep_a), the second is equal to 1 when it is a type B firm (bvd_indep_b) and the third is equal to one when it is a type C firm (bvd_indep_c). Reference category is type U firms. See text for more details.
Controls	Macroeconomic environment: Year dummies are included. The reference year is 2000.
	Sector affiliation dummies: Dummies for the following sectors are included (see section 14 (appendix)): Farming, Forestry, Fishing, Mining, Manufacturing (reference dummy), Energy, Construction, Trade and hotel, Transport, Business service, Public service activities, Organisations. The following abbreviations are used. Dumfar = Farming, forestry and fishing, Dummin = Mining, Dumman= Manufacturing, Dumener = Energy, Dumcon = Construction, Dumtraho = Trade and hotel, Dumtra = Transport, Dumbus = Business service, Dumpub = Public service activities, Dumorg = Organisations. As there are no NACE codes for the IT and tele-sector a (self-constructed) IT and tele-dummy is included in the estimations. On top of belonging to one of the above sectors a firm is considered to be in the IT and tele group if it has activities in one of the sectors listed in table 14.b in section 14 (appendix).

Hypothesis 7: Legal form is used as a proxy for willingness to take on risk. The legal form dummy is constructed such that for all countries, the dummy is equal to one, when the legal form of the company is a private limited liability company, and it is equal to zero, when it is a public limited liability company. One would think that private limited liability companies would be more risky, as they would have less share capital to loose compared with public limited liability companies, and so the variable is hypothesized to have a positive sign.[13]

[13] The terms public and private limited liability companies are used to denote the law the specific company is following. There are different rules for public and private limited liability companies, i.e. concerning the amount of share capital that they need to hold. Note that not all public limited liability companies are listed on a stock exchange.

Hypothesis 8: Ownership information is included in the estimations as a proxy for the firms' internal environment. The governance of a firm, and so its financial decisions, is influenced by the ownership structure of the firm. BvD provides an independence indicator measuring the degree of independence of a company (management) with regard to its shareholders. It is equal to 1) A, when none of the company's shareholders has more than 24.9 pct. of ownership share, 2) B, when none of the shareholders have an ownership percentage over 49.9 pct., but at least one or more shareholders has an ownership percentage above 24.9 pct., 3) C, when at least one of the shareholders has an ownership over 49.9 pct. and 4) U, when there is no information on the shareholders. In the estimations, not having an independence indicator is the reference category. Three dummies are then included, one for type A firms (measured by bvd_indep_a), one for type B firms (measured by bvd_indep_b) and one for type C firms (measured by bvd_indep_c). Analysis of the potential conflict between owners leads to the result (among other results), that it is desirable to concentrate ownership among few individuals.[14] The hypothesis is therefore, that the bvd_indep_c variable has a negative sign. In this way, firms that are of type C are thought to be more concentrated than the reference category. No hypotheses are set up on the effect of bvd_indep_a and bvd_indep_b. It is difficult to say something about these types of firms relative to the reference category. Compared to firms of type C, the effect of being a type A or a type B firm is expected to have less effect.

Hypothesis 9: On top of the independence indicator, a variable, which measures the number of shareholders in the firms, are included. This variable measures the number of shareholders and not the degree of independence, however, it is correlated with the independence indicator (as the ownership of firms with many shareholders are more likely to be less concentrated). Hermalin and Weisbach (2003) discuss the agency problem between the owners of a firm (its shareholders) and the management. According to their introduction[15], it seems preferable to have a large outside shareholder. This would be along the lines of Bennedsen and Wolfenzohn (2000), who argue, that it is desirable to concentrate ownership among few individuals, and so the hypothesis may be that a large number of shareholders increase the likelihood of financial distress (despite the fact that the holdings of the shareholders cannot be observed in the data). But there may also be other factors at play. Firms have boards, because they

[14] See Bennedsen and Wolfenzon (2000). In the academic literature, broadly, two different types of ownership for limited companies have been discussed, c.f. Bennedsen (2004), who provides an overview of the literature. One type is the Anglo-Saxon business structure, which is dominated by many small owners and strong management. The other type is the Continental European Business structure, where there are controlling shareholders who often participate in everyday management of the company. Bennedsen (2004) explains: "The most important management problem is … whether there are conflicts – or the potential for conflicts – between the controlling majority shareholders and the non-controlling minority shareholders. … There are … many examples of how … corporations have been wrecked because of disagreements and the resulting conflicts between owners".

[15] Hermalin and Weisbach (2003:10): Managers "tend to be insufficiently vigilant or trustworthy … One solution to this problem is to provide management with strong incentives contractually. But this begs the question of who provides these incentives and who ensures that the incentive contracts are structured optimally? In most large corporations, the shareholders are too diffuse, rationally plagued by a free-rider problem, and, for the same reason, too uninformed to set managers' compensation. This problem, as well as the underlying direct control problem, could be alleviated in situations in which a large outside shareholder has sufficient incentive herself to tackle them. … While there are certainly instances in which large shareholders play an important governance role, this is also certainly not a universal solution".

are "part of the equilibrium solution to the contracting problem between diffuse shareholders and management", (Hermalin and Weisbach (2003:10)). As there is no information on boards in the dataset and no clear hypothesis can be set up on the effect on the number of shareholders, it will be left to the estimations to show the effect of this variable.

Sector affiliation dummies are included to control for the sector affiliation of the firms. The reference category is firms in the manufacturing sector. Furthermore, year dummies are included to control for business cycle effects. The reference year is 2000.

3.3 Sample selection

The construction of the sample used in the estimations is discussed in this section. The sample selection criteria are summarized in table 3.3, which is split up on conceptual criteria and other criteria.[16] After the application of the criteria, the database consists of a total of 282,131 firm-year observations (FR: 108,533, IT: 97,732, ES: 75,866) covering the years 2000 - 2002. The proportion of SMEs in financial distress to the overall number of SME firm-years is 0.2 per cent in Spain and Italy and 1.6 per cent in the French case.[17] Despite the differences in levels, it is the assessment, that we have a random sample, and accordingly, that the estimations are consistent across countries and that the effects of the explanatory variables in the various countries can be meaningfully compared. Descriptive statistics are found in section 13 (appendix).

Most of the sample selection criteria are self-explanatory. Nonetheless, in the following sections, it is discussed why only public and private limited liability companies have been considered and why it was chosen to focus only on SMEs. Furthermore, the attrition problem is discussed.

[16] See section 11 (appendix) for details.

[17] To benchmark the French, Spanish and Italian data, they are compared to a sample of Danish SME's, which covers the whole population of Danish public and private limited liability companies. The sample is analysed and discussed extensively in Dyrberg (2004). In this sample of Danish firms, the proportion of E1 events is 0.8 per cent, which is higher than the fraction in Italy and Spain and lower than fraction in France. Compared to the dataset used in Dyrberg (2004) the following corrections are made in order to make the figures comparable: Only SMEs are considered and the E1 measure is modified to be comparable to the E1 measure used in this paper. The Danish data set includes bankruptcies in the period 1995 – 2001. The dataset for the other countries covers the period 2000 – 2002.

Table 3.3: Sample selection criteria

Criteria	
Conceptual	Only unconsolidated statements are analysed.
	Financial institutions and non-financial holding companies are excluded.
	Only public limited liabilities and private limited liabilities are analysed.
	Only SMEs are analysed.
	Some firms leave with no explanation (that is, they are not assigned an exit code). These firms are called attritioners. They are denoted E5. They are excluded from the estimations based on bvd_id.
Other	If a company hands in two financial statements in one year, only the last financial statement is included in the estimations.
	Active companies are excluded if they hand in a financial statement in 2000 and 2002 only.
	Various corrections are made to the database (e.g. firms with illogical variables, such as short-term debt less than zero and a solvency ratio larger than 100 pct., are excluded).
	Firms with missing variables on any of the explanatory variables are excluded.

3.3.1 Public and private limited liability companies and SMEs

In this section it is discussed how it is ensured that we have a homogenous group of firms across countries.

A study on bankruptcies and the legal consequences of bankruptcies is provided by the European Commission (2003a). The work takes its point of departure in the Principles and Guidelines for Effective Insolvency and Creditor Rights Systems, which were developed by the World Bank to promote an international consensus on a framework to assess the effectiveness of insolvency and creditor rights systems.[18] Attached to the European Commission (2003a) study are country studies, which among other things include detailed discussions of the accounting standards in the countries, c.f. European Commission (2003b, 2003c and 2003d). These discussions are reviewed in this section. The overall conclusion is that there are many nuances in the countries, but that it makes sense to distinguish a homogenous group of firms across countries, namely SMEs that are either private or public limited liability companies.

[18] The European Commission set up an expert group in 2002 to take part in a benchmarking exercise chaired by the Commission. It consisted of experts from 14 Member States, 7 Candidate Countries and Norway. The European Commission (2003a:321) points out that the "results of the … questionnaires that we received from our experts are to be taken carefully and to be considered as nothing more than what they really reflect: the opinion of … national experts regarding the implementation of the World Bank principles in their own legal systems, based on their high experience in the matter of insolvency. Accordingly, we believe that it is interesting to show and to describe practices throughout the EU Member States and the U.S. regarding the World Bank principles, as they are perceived by the national experts. Nevertheless, we are aware that their results cannot necessarily be extended or generalized, and that is the reason why we would not affirm Member States that have the highest rate of implementation should be showed as examples of best practice." According to the study, out of the three analysed studies, Spain has the largest number of principles not adopted. France and Italy come out similarly.

In France, public limited liability companies ("société par action"[19]), private limited liability companies ("société à responsabilité limitée") and partnerships ("société en nom collectif") must file annual accounts and an annual report with the clerk of the Commercial Court and they are subject to criminal penalties in case of violation (European Commission (2003d:3f)).

In Italy, public limited liability companies ("Società par Azoni") and private limited liability companies ("Società a Responsabilità Limitata"), whose share capital is equal to or higher than 100,000 euro or companies who for two years has not provided the duly publication of the balance sheets required by law, are supervised by a board of internal auditors. The board has the task to control the management of the company in order to safeguard the interest of the shareholders and creditors of the company. It is "… constituted by professionals registered with the Roll of Certified Accountants and appointed by the shareholders, shall supervise the management of the company, the compliance by the other corporate bodies with applicable legal and statutory rules and it controls that the company's accounts are regularly kept, that the balance sheet reflects the situation resulting from the company's accountancy books and that the rules established for the evaluation of the company's assets are complied with", c.f. European Commission (2003b:7).

In Spain, the filing is mandatory for public limited liability companies ("Sociedad Anonima") and private limited liability companies ("Sociedad Limitada").[20] Despite the rules, there are quite a number of companies, which do not comply with the annual accounts record obligation (European Commission (2003c:10)). One could fear that the weaker firms are the ones that are not handing in their financial statement. If this is true, the estimates for Spain would be conservative compared to the actual situation. Having this in mind, one could choose to focus on audited firms only. In Spain, firms whose turnover exceeds 4.75 million euros, whose total assets exceed 2.37 million euros and whose total number of workers exceeds 50 during two consecutive years, are obliged to audit their financial statements.[21]

Table 3.3.1: The definition of micro, small and medium-sized enterprises

Category	Employees	Turnover OR Total balance sheet
Micro enterprises	<10	<= 2 million euro OR <= 2 million euro
Micro and Small enterprises	<50	<= 10 million euro OR <= 10 million euro
Micro, Small and Medium-sized enterprises	<250	<= 50 million euro OR <= 43 million euro

Note: A further criterion for being an SME according to the European Commission definition concerns the economic power of the enterprise. According to the European Commission a distinction should be made between various types of enterprises, depending on whether they are autonomous, whether they have holdings, which do not entail a controlling position or whether they are linked to other enterprises. For an SME to be considered autonomous, less than 25 pct. capital shares should be held by third party. This criterion is not taken into account in the construction of the SME sample. Instead an independence indicator is included in the estimations. Source: European Commission (2003e).

[19] In the French country study, the expression "Société par action" is used to denote public limited liability companies. "Société par action" is not a legal form itself, but comprises of a set of companies, including "société anonyme à conseil d'administration", "société anonyme à directoire" and "société par actions simplifiée".

[20] Information is provided from Informa S.A. in October 2004. Informa S.A. provides Spanish data to Amadeus.

[21] We would like to thank Antonio Marcelo for pointing this out and for providing us with the information. The information is as of September 2004.

To address the potential problem with the financial statements from Spanish firms, it was decided to follow the European Commission definition of SMEs (see table 3.3.1) and to stick to this definition for all countries. According to the SME lower bound criterion firms with at least 10 employees and with total assets of at least 2 million euro are included in the sample.[22] It is not as strict as the criterion for when Spanish firms should be audited. Nonetheless, the criterion does accommodate the critique posed by Creditreform. According to Creditreform (2003:16 and 2002:8), micro-companies cannot file for bankruptcy in Spain, and companies, which can no longer pay their bills, are not brought to court by their creditors, as the relevant proceedings are too elaborate and too costly to the creditors to justify the effort involved. Creditreform (2002:8) notes: "To permit a real comparison with, say, Germany, all those cases rejected in this country for lack of sufficient assets to justify proceedings would have to be omitted from the statistics." The criterion ensures that micro-companies, which resemble households, are excluded from the sample, and furthermore, that only "truly" active companies are analysed. The upper bound criterion ensures that the analysed group of firms is fairly homogenous.

To align the analysed group of French companies with the group of Spanish and Italian companies, which do not include partnerships, only French, Italian and Spanish public limited liability companies and private limited liability companies are analysed.

3.3.2 Attrition

The fact that some firms leave with no explanation (that is, they are not assigned an exit code), is called attrition. It is crucial to find out whether or not attrition is a problem. If there is a selection bias, i.e. a distortion of the estimation results due to non-random patterns of attrition, we need to correct for this.[23]

This section starts out discussing the extent of the attrition problem and by explaining why the drop-outs could potentially be a problem for the analysis. Afterwards the various types of drop-outs discussed in the literature are briefly reviewed.

---

[22] The criteria are applied to the firms the year they are included in the database. This means, that the position of the firms can be deteriorating once they are included.

[23] Attrition is a common problem seen in many panel data studies. Accordingly how to treat attrition in panel data models is discussed in a number of articles, c.f. the discussion of attrition in Wooldridge (2002:585ff) and the Journal of Human Resources (Spring 1998), which has a special issue devoted to the topic. Alderman, Behrman, Kohler, Maluccio and Watkins (2000) summarize the conclusion from the papers in The Journal of Human Resources with the words: "The striking result of these studies is that the biases in estimated socioeconomic relations due to attrition are small – despite attrition rates as high as 50 percent and with significant differences between the means of a number of outcome and standard control variables". This is an interesting conclusion that they themselves also obtain in their own study, in which they discuss the extent and implications of attrition in three longitudinal household surveys from Bolivia, Kenya and South Africa, which all report very high per-year attrition rates between survey rounds. The attrition rates are considerable: 35 percent for the Bolivian sample, between 28 percent for women to 41 percent for couples in the Kenyan sample and form 16 percent for households to 22 percent for preschool children in the South African sample (Alderman et al. (2000:12)). The conclusion in Alderman et al. (2000:24) is that "…in contrast to often-expressed concerns about attrition, for many estimates the coefficients on standards variables in equations are unaffected by attrition … Thus, even when attrition is fairly high, as it is in the samples we used, attrition apparently is not a general and pervasive problem for obtaining consistent estimates.". This being said, it is stressed in the paper, that, as a general observation, analysts should assess the problem for the particular model and the particular data that is used.

The extent of the attrition problem in the dataset is seen from tables 3.3.2.a, 3.3.2.b and 3.3.2.c. The tables give an overview of the data, when all sample selection criteria, except the attrition sample selection criteria, are taken into account (see table 3.3 in section 3.3). Table 3.3.2.a shows, that in 2000, 73 Italian firms left the database because of an E1 event, 6 firms left the database because of an E2 event, 117 firms left the database because of an E3 event etc. Throughout the whole period (2000 – 2002), a total of 218 Italian firms left the dataset because of financial distress (E1 event), 24 firms left because of mergers (E2 event), 439 firms left because they are voluntarily liquidated (E3 event) etc. It is important to note that 5,201 Italian firms are recorded as E5 firms (total). 1,278 Italian firms are "lost" between 2000 and 2001, and so they are recorded as E5 firms in 2000, and 3,923 firms are "lost" between 2001 and 2002, and so they are recorded as E5 firms in 2001. The E5 firms belong to the category "unknown whether the firm has exited or if it is still active". The only thing, which is known about these firms, is that they are no longer observed in the database. The firms that are recorded as E5 firms in year 2000 are no longer observed in the years 2001 and 2002. The firms that are recorded as E5 firms in year 2001 are no longer observed in year 2002. The recording of E5 firms is zero for 2002, as no firm is known to be lost after 2002 (as the dataset stops in 2002). The E5 firms are called attritioners. The fact that they leave the panel (with no explanation) is called attrition.

Table 3.3.2.a: Italy: E1, E2, E3, E4, E5 and active firms

	2000	2001	2002	Total
E1	73	72	10	155
E2	6	2	16	24
E3	117	74	248	439
E4	9	10	46	65
E5	1278	3923	0	5201
Active	31316	32989	35977	100282
Total	32799	37070	36297	106166

Table 3.3.2.b: Spain: E1, E2, E3, E4, E5 and active firms

	2000	2001	2002	Total
E1	65	64	51	180
E2	31	41	23	95
E3	272	301	344	917
E4	8	15	27	50
E5	519	2204	0	2723
Active	21586	25260	29349	76195
Total	22481	27885	29794	80160

Table 3.3.2.c: France: E1, E2, E3, E4, E5 and active firms

	2000	2001	2002	Total
E1	167	396	1140	1703
E2	48	686	675	1409
E3	2	22	39	63
E4	152	418	525	1095
E5	873	2490	0	3364
Active	30281	34710	40796	105787
Total	31523	38722	43175	113420

One could have two opposing hypotheses of why some firms drop-out for unknown reasons (the E5 firms). *Hypothesis 1*: The first hypothesis is that Bureau van Dijk (BvD) does not care about the firms, when they are no longer active, and therefore that it is not sure that all non-active firms get an exit code. Note, however, that if BvD want to track the bankrupt firms, it should not be difficult, as the names of the firms that have gone bankrupt are usually published. *Hypothesis 2*: The other hypothesis is that the firms that drop out for unknown reasons are likely to be active firms, as firms that exit because of financial distress are easier to track than active firms. For statisticians it can be a really time-consuming task to trace active firms that change legal form or for other reasons re-register, c.f. Statistics Denmark (2002) and Eurostat (2004). If the active firms in Spain, Italy and France change legal status or for other reasons are registered in another way, and if they are not traced, this may explain part of the attrition observed in the data.

A priori it is not clear which of the hypotheses are true, or if it is a mix of them. If Hypothesis 1 is true, then one would think that the proportion of firms in financial distress is larger among the drop-outs, than in the sample as a whole. If Hypothesis 2 is true, then one may think that the proportion of firms in financial distress is likely to be the same among drop-outs and non-drop-outs, or maybe that the firms in financial distress is under-represented in the drop-out group. The hypothesis about the drop-outs affects the estimation procedure and it is therefore important to find out what type of drop-outs one deals with. If there is a selection bias, i.e. a distortion of the estimation results due to non-random patterns of attrition, one needs to correct for this.

To formalise the discussion, the various types of drop-outs discussed in the literature are briefly reviewed. The next sections draws on Diggle and Kenward (1994), Fitzmaurice, Heath and Clifford (1996), Fitzgerald, Gottschalk and Moffitt (1998) and Feelders (2003).

The main distinction to make is between selection on observables and selection on unobservables. We talk about selection on observables, when firms drop out 1) completely at random or 2) randomly. A firm drops out *completely at random*, if firms with low earnings are just as likely to be in the sample as firms with higher earnings, or if firms with a high solvency ratio are just as likely to be in the sample as firms with a low solvency ratio. The definition of firms that *drop out randomly* is a little different. In

the literature random drop outs are used for firms that are better described as firms that drop out randomly within a class. Here is an example. If small firms are less inclined to report their income, then reported income will be related to the size of a firm. If, within small firms, the probability of reported income is unrelated to the income level, then the data would be considered missing at random, but it would not be considered missing completely at random. When firms drop out completely at random or drop out at randomly, the firms can be ignored in the estimations.

Selection on unobservables is the term, which is used when the drop-out process is informative, and the *drop-outs are non-ignorable*. In this case, the drop-outs depend on unobserved factors or, in other words, something which cannot be measured. Examples of non-ignorable drop-outs would be if firms with incompetent managers (incompetence is something which is inherently unobservable) are more likely to be among the drop-outs, or if the firms that drop-out, are firms facing a greater degree of uncertainty (something which is inherently unobservable also). If the drop-outs are non-ignorable they cannot be excluded from the estimations. Depending on the assumptions made concerning the drop-outs, the estimation problem can become very difficult to handle.

The likelihood of selection on unobservables is smaller, the more variables there is included in the estimations, as there is less unobservable variation left. A number of proxy variables are used in this paper for inherently unobservable variables, and so selection on unobservables will not be investigated further. Instead the focus will be on selection on observables. The approach of Alderman, Behrman, Kohler, Maluccio and Watkins (2000) is followed. Tests of equal means of the explanatory variables are undertaken, attrition probits are estimated, and, as a robustness check, the credit-scoring models are estimated using 1) a dataset with the drop-outs (the E5 firms) as an exit option even though we do not know if they are "real" exits and using 2) a dataset without the E5 firms and accordingly, with no E5 exit option, and, thereafter, the results are compared.[24] The detailed results are reported in section 12 (appendix).

The comparisons of the results from the attrition probits and the comparison of means show the following: For France, the comparison of the attrition probits with the results obtained in the sections on comparisons of means shows that the results are very alike. The overall conclusion using both tests is that the variables, which are central to this study, do not differ in a systematic way between E5 firms and non-E5 firms. For Italy, the results show, that the characteristics, which predict attrition with multivariate controls, and what the directions of those effects are, cannot be inferred simply by examining the significance of means in univariate comparisons between the subsamples. The two

[24] This robustness check differs from the approach in Alderman et al. (2000). As the selection in Alderman et al. (2000) is not on the dependent variable, in their paper, it is tested whether coefficient estimates differ for two subsamples. They compare the results using the total sample and the results using the nonattriting sample. The idea is that the parameter estimates of the total sample would be different from the parameters estimates using the nonattriting sample, if the attritioners are different from the nonattritioners. Here this test cannot be done. Instead, as a robustness check, the credit-scoring models are estimated using 1) a dataset with the E5 firms (and with the E5 exit as an exit option) and using 2) a dataset without the E5 firms (and accordingly, with no E5 exit option), and, thereafter, the results are compared (significance and sign of parameter estimates as well as predictive ability).

methods lead to opposing results: While the comparisons of means suggested that worse-off firms may be more likely to be among the attritioners, the multivariate estimates are less supportive of this conclusion. For Spain, the results go both ways. There are indications that firms that are worse-off are among the attritioners (based on the solvency and the earnings ratio). On the other hand the results on the loans to assets ratio indicate that the E5 firms are better off. The results coming from the comparisons of the means and the attrition probits on age, size and the proxies are conflicting, indicating that these variables show no clear pattern in the potential bias of the E5 firms.

The robustness checks show that the credit-scoring models estimated using 1) a dataset with the E5 firms and with the E5 exit as an exit option, even though we do not know if they are "real" exits, and using 2) a dataset without the E5 firms and accordingly, with no E5 exit option, perform similarly, when their predictive ability is compared. Furthermore the same variables are significant using the two specifications (and with the same signs), except one variable for Italy.

Based on the analysis, the overall conclusion is that the exclusion of E5 firms does not seem to bias estimates, and so they are excluded from the dataset (based on their bvd_id).

4. Econometric Theory

This section presents the estimation method. As firms can exit for other reasons than financial distress, e.g. they can merge with other firms, the credit-scoring models are estimated as competing-risks models. A competing-risks model estimates the probability of exiting to various states. Here the relevant states are E1 (financial distress), E2 (merger), E3 (voluntary liquidation) and E4 (inactive (no precision)). The E1 hazard, $h_{E1}(t)$, measures the probability of exiting to financial distress at time *t* given that the firm made it to the current period, i.e. the probability that a firm exits as an E1 firm in 2002 given that the firm made it to 2002. The E2 hazard, $h_{E2}(t)$, measures the probability of being merged at time *t* given that the firm made it to the current period etc. Depending on the data at hand (continuous data, grouped duration data or discrete duration data) and the assumptions one is willing to make, various strategies for the estimation of a competing-risks problem exist. Jenkins (2003) provides a detailed technical discussion of the main assumptions made in the literature and the implications of the various assumptions. For this estimation problem we have grouped duration data. It is chosen to treat data as intrinsically discrete and to follow the estimation strategy of Allison (1982). First the exits are assumed to be independent. This implies that the likelihood function can be written as:

$$L = (L_{E1})^{\delta_{E1}} (L_{E2})^{\delta_{E2}} (L_{E3})^{\delta_{E3}} (L_{E4})^{\delta_{E4}} (L_{active})^{1-\delta_{E1}-\delta_{E2}-\delta_{E3}-\delta_{E4}}$$

$$= \left[\frac{h_{E1}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)-h_{E4}(t)}\right]^{\delta_{E1}} \left[\frac{h_{E2}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)-h_{E4}(t)}\right]^{\delta_{E2}}$$

$$\times \left[\frac{h_{E3}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)-h_{E4}(t)}\right]^{\delta_{E3}} \left[\frac{h_{E4}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)-h_{E4}(t)}\right]^{\delta_{E4}}$$

$$\times \prod_{\tau=1}^{t} [1-h_{E1}(\tau)-h_{E2}(\tau)-h_{E3}(\tau)-h_{E4}(\tau)]$$

where τ denotes the year the firm gets incorporated (this of course differs across firms), and where $\delta_{E1} = 1$ when the specific firm exits because of E1, $\delta_{E2} = 1$ when the specific firm exits because of E2 etc. When the firm does not exit for E1, E2 reasons etc., it is active (and gets censored in 2002).

The likelihood function cannot be factored into separate components and so maximum likelihood estimation must be done simultaneously for all kinds of events. Allison (1982) shows that the estimation problem becomes very easy to handle if one assumes a particular form of destination-specific hazards. By assuming that each of the destination-specific hazards have the following form (which, of course, has to be modified for the other exit possibilities)

$$h_{E1}(t) = \frac{\exp(\beta'_{E1} X_t)}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)},$$

where

X_t characterizes the covariates (including age, i.e. the baseline-hazard function. The baseline-hazard function is specified non-parametrically, as dummies for each age are included. Note that the age of a firm differs from calendar time),

β are the parameters of the covariates,

the likelihood function can be rewritten to have the same form as a standard multinomial logit model, and so it can be estimated by standard methods. This result is obtained when the destination-specific hazard functions are substituted into the likelihood function from above.

Inserting the destination-specific hazard functions in the likelihood functions then gives

$$L = \left[\frac{\exp(\beta'_{E1} X_t)}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)}\right]^{\delta_{E1}}$$

$$\times \left[\frac{\exp(\beta'_{E2} X_t)}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)}\right]^{\delta_{E2}}$$

$$\times \left[\frac{\exp(\beta'_{E3} X_t)}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)}\right]^{\delta_{E3}}$$

$$\times \left[\frac{\exp(\beta'_{E4} X_t)}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)}\right]^{\delta_{E4}}$$

$$\times \left[\frac{1}{1 + \exp(\beta'_{E1} X_t) + \exp(\beta'_{E2} X_t) + \exp(\beta'_{E3} X_t) + \exp(\beta'_{E4} X_t)}\right]^{1-\delta_{E1}-\delta_{E2}-\delta_{E3}-\delta_{E4}}$$

$$\times \prod_{\tau=1}^{t-1} \frac{1}{1 + \exp(\beta'_{E1} X_\tau) + \exp(\beta'_{E2} X_\tau) + \exp(\beta'_{E3} X_\tau) + \exp(\beta'_{E4} X_\tau)}$$

which is exactly the same as the likelihood function for the multinomial logit model. The estimation problem can therefore easily be estimated. Active firms are chosen as the reference category. Left truncation and right censoring is handled as in Jenkins (1995), Henley (1998:418) and Rommer (2005).

The coefficients that are reported from the estimation of the problem must be interpreted as contrasts between pairs of categories. As the focus in this study is on finding the effects of the E1 firms, the relevant equation to analyse is the E1 hazard:

$$\log(\frac{h_{E1}}{h_{active}}) = \beta'_{E1} X_t$$

The interpretation of the equation is that if, for example, the parameter estimate on the solvency ratio is -1.5832, then each 1-level increase in the variable multiplies the odds of moving into financial distress versus staying active by about 0.21 (= exp(-1.5832)).

This above equation is obtained for France, Italy and Spain. In section 5 the parameter estimates are compared and discussed.

Note that from the estimations one also obtains the parameter estimates for the other exits stemming from the following equations:

$$\log(\frac{h_{E2}}{h_{active}}) = \beta'_{E2} X_t$$

$$\log(\frac{h_{E3}}{h_{active}}) = \beta'_{E3} X_t$$

$$\log(\frac{h_{E4}}{h_{active}}) = \beta'_{E4} X_t$$

These results are reported in section 15 (appendix), however, as they are not the focus of this paper, there are not discussed and interpreted.

As a last point, note that unobserved heterogeneity is not allowed for in the estimation problem. This means that if two firms have identical values on the covariates, they also have identical hazard functions, or, in other words, all differences between firms are assumed to be captured using observed explanatory variables. Proxy variables as well as a flexible baseline-hazard specification are used in the estimations. This should mitigate the effects of unobserved heterogeneity, c.f. the discussion Jenkins (2003:102).

5. Results: The Country Models

5.1 Parameter Estimates

Section 15 (appendix) presents the parameter estimates on the E1, E2, E3 and E4 hazard in the three countries. The expected sign and the estimated signs and the significance levels of the E1 hazards are

summarized in table 5.1.a (core variables) and table 5.1.b (proxies). Table 6.1 in section 6.1 depicts the parameter estimates of the E1 hazards in the three countries as well as in the pooled model.

Table 5.1.a: Core variables

Variable	Expected sign			Estimated sign		
	Italy	France	Spain	Italy	France	Spain
Profitability: Earnings ratio	-	-	-	-	-	-
Solvency: Solvency ratio	-	-	-	-	-	-
Leverage: Loans ratio	+	+	+	Insign.	+	+
Firms size: Log(total assets)	-	-	-	+	-	Insign.
Age	-	-	-	-	+	Insign.

Note: Expected sign: - denotes that the variable is expected to affect financial distress negatively. + denotes that the variable is expected to affect financial distress positively. ? denotes that no certain effect is expected. Estimated sign: - denotes that the variable does affect financial distress negatively. + denotes that the variable does affect financial distress positively. Insign. denotes that the variable is insignificant in the estimations. A significance level of 5 pct. is chosen. Year dummies and sector affiliation dummies were included. Because the data was too sparse otherwise in some countries, a grouping of the sector affiliation dummies took place. France: None of the age dummies were significant in the first estimations. In the final estimations no age dummies are included, only the variable age. Italy and Spain: In the first estimations a flexible baseline-hazard function was specified. This led to a quasi-complete separation of data points, meaning that a maximum likelihood estimate may not be possible to obtain, as the data is too sparse. The consequence of this was to use age in the estimations.

Table 5.1.b: Proxies

Variable	Expected sign			Estimated sign		
	Italy	France	Spain	Italy	France	Spain
Subsidiaries: The number of subsidiaries that a company has registered	?	?	?	Insign.	Insign.	Insign.
Legal form: The effect of being a private limited liability company	+	+	+	+	Insign.	Insign.
Shareholders: The number of recorded shareholders	?	?	?	Insign.	Insign.	Insign.
Ownership variables:						
Bvd_indep_a	?	?	?	Insign.	Insign.	Insign.
Bvd_indep_b	?	?	?	Insign.	Insign.	Insign.
Bvd_indep_c	-	-	-	Insign.	-	Insign.

Note: For further details see the note to table 5.1.a.

The tables show that there are some similarities across countries, but also that there are quite a lot of differences between the countries. The determinants of financial distress that behaves similarly across countries are 2 of the core variables, namely the earnings ratio and the solvency ratio (they are significant and have a negative sign in all countries, indicating that the higher these ratios are, the less likely a firm is to enter financial distress), and 4 of the proxy variables, namely the number of subsidiaries and shareholders, the ownership variables bvd_indep_a and bvd_indep_b (they are insignificant in all countries).

The variables, which differ between the countries in terms of whether or not they are significant or what sign they have, are the loans to total assets ratio, size, age, legal form and bvd_indep_c.

Loans to total assets were assumed to have a positive coefficient. The parameter estimate is significant and has a positive sign in the French and the Spanish case. The variable is insignificant in the Italian credit-scoring model. A reason for this could be that in comparison with French and Spanish firms, Italian firms fund themselves to a greater extent through trade creditors, c.f. European Committee of Central Balance Sheet Offices (2000:20). Another explanation, which is not mutually exclusive with the first explanation, could be that despite the common feature of "multi-banking" (i.e. firms maintain relations with a number of banks) in all countries, the use of multiple credit lines has become especially large in Italy, c.f. European Committee of Central Balance Sheet Offices (2000:36).[25] According to the European Committee of Central Balance Sheet Offices (2000:36f), in Italy, "… the bank bases its lending decisions on real or personal guarantees enabling it to keep its risk exposure within acceptable limits. Moreover, it is much easier and quicker to assess guarantees than to evaluate the company: in the first case, all one needs to do is to evaluate a specific asset (property or financial) and the quality of a surety signatory, making it easier from an administrative viewpoint to provide the specific economic-legal professional training and to process the loan application. The bank therefore bases its lending policy on three closely-related components: spreading the risk by limiting the amount of credit granted to each customer; lending using technical methods allowing for the immediate revocation of the loan application; and the existence of guarantees or surety to cover losses arising from insolvency. This model, which has developed in Italy partly due to a legal framework for insolvency and bankruptcy that systematically gives priority to secured creditors ahead of other categories, limits the bankruptcy risk borne by banks."

Age and size were hypothesized to affect the firms in financial distress negatively. The estimations show that these variables are insignificant in the Spanish case, but significant in the Italian and French case. As mentioned in section 3.2 the size and the age variables are correlated, and so it can be difficult to disentangle the effects. This is what is seen from the estimations. In France and Italy both variables

[25] The phenomenon of "multi-banking" is described in European Committee of Central Balance Sheet Offices (2000:36): "A firm … obtains financing by placing banks in competition with each other to obtain the most favourable contractual terms regarding interest rates, services and maturities. The bank spreads its risks by having smaller commitments and can recover its money inasmuch as the firm can take another loan from another bank".

are significant, but with changing signs so that if age is significant and has a positive effect in one country, then size is significant and has a negative effect in the same country, or vice versa. In the French case age affects the likelihood of entering financial distress positively, and size affects the likelihood of entering financial distress negatively. The least significant variable is age. If age is left out of the estimations, size is significant and has a negative sign. In the Italian case the least significant variable of age and size is size. If size is left out of the estimations, then age is negative and significant.

The legal form dummy is constructed such that for all countries, the dummy is equal to one, when the legal form of the company is a private limited liability company. The legal form variable was hypothesized to have a positive sign. The results show that the variable is only significant in the Italian case, where it does have the hypothesized positive sign. The level of share capital between public and private limited liability companies differs between the countries. In Italy the difference in share capital between the two types of legal forms is 110,000 euro, in Spain it is 60,000 euro and in France it is 37,000 euro[26]. As only firms with 10 employees and a balance sheet of at least 2 million euro are considered in the estimations (c.f. section 3.3), it is not surprising that only an effect of the private limited liability variable for the Italian firms, for which the difference in share capital between the private and public limited liability companies is the largest, is significant.

Ownership information is also included in the estimations. The hypothesis is that the bvd_indep_c has a negative sign. The estimations show that there is no effect for the Italian and Spanish firms and that the parameter estimate is significant for the French firms and has a negative sign. The reason for this result could be that the French firms are by far the most concentrated firms, c.f. the descriptive statistics in section 13 (appendix).

5.2 Discriminatory Power

A measure of how well the model with the specifications for the hazard function fits the data is the proportion of correct predictions. There is a trade-off between incorrectly classifying a firm that does

[26] This argument does not take the few French firms in the sample that are listed into account. Different rules apply to the different countries concerning the amount of share capital that the firms need to hold. In Spain, public limited liability companies need to hold the minimum share capital of around 60,000 euro, whereas private limited liability companies do not need to hold minimum share capital (Information is provided by Informa S.A. in October 2004. Informa S.A. provides Spanish data to Amadeus). In Italy, public limited liability companies need to hold at least 120,000 euro as share capital, and private limited liability companies need to hold at least 10,000 euro (These new rules came into force in January 2004, c.f. Cappiello and Marano (2003:3). Before then public limited liability companies needed to hold 100,000 euro and private limited liability companies needed to hold 10,000 euro, c.f. European Commission (2003b:6)). In France, public limited liability companies need to hold 37,000 euro as share capital (except the ones that are listed on a stock exchange), whereas limited liability companies do not need to hold any share capital (The information on the French firms is as of November 2004. Public limited liability companies: "Société anonyme à conseil d'administration" needs to hold 37,000 euro as share capital. If they are listed on a stock exchange, called "appel public à l'épargne", they need to hold at least 225,000 euro. (Broadly, "appel public à l'épargne" corresponds to being listed on a stock exchange. However, it is not exactly the same, as "appel public à l'épargne" is a broader concept. It means "selling shares to the public"). "Société anonyme à directoire" need to hold 37,000 euro as share capital. They distinguish themselves from "société anonyme à conseil d'administration" in the way they are governed. "Société par actions simplifiée" need to hold 37,000 euro as share capital. They do not have the possibility of "appel public à l'épargne". Private limited liability companies: As of August 1st, 2003, "Société à responsabilité limitée" do not need to hold any share capital. Before August 1st, 2003, they needed to hold a minimum of 7,500 euro as share capital).

not exit because of financial distress as a financially distressed firm compared to not classifying a financially distressed firm as financially distressed.

The naïve predictor uses a cut-off value of 0.5, which means that firms that have a predicted probability above 0.5 are classified as financially distressed firms, and that firms that have a predicted probability below 0.5 are classified as active firms. A cut-off level of 0.5 would have seemed reasonable if the samples had entailed 50 per cent financially distressed firms and 50 per cent active firms (or other firms that were not financially distressed). In that case the ratio of financially distressed firms to all other firms would have been exactly 0.5. As the samples sizes here are skewed with only a small fraction of firms in financial distress compared to all other firms in all countries, the 0.5 cut-off is modified. As a cut-off level in the country models the proportion of financially distressed firms used in the estimations to all other firms is used.[27] Tables 5.2.a, 5.2.b and 5.2.c. and the more detailed tables 15.d, 15.e and 15.f in section 15 (appendix) show the number of correct predictions, the number of correctly called non-events, as well as the number of type I errors (missing prediction) and type II errors (wrong signal) in the three countries. With the used cut-off levels the models correctly classify between 75 and 88 per cent of the financially distressed firms as financially distressed, and they correctly classify between 68 and 72 per cent of the active firms as active. Had one chosen to use a lower cut-off level, one would have predicted more firms to be financially distressed, but this would be at the cost of an increased number of type 2 errors (wrong signal). If the cut-off level is increased, one would make less type 2 errors, but that would be at the cost of a decreased number of firms that are predicted to be in financial distress (and an increase in the type 1 errors).

The policy maker or the credit institution that uses the model has to make the decision on what cut-off level to use. The cut-off level depends on the "agents" objective function, also called the loss function. It should reflect an assessment of the cost of making type I and type II errors, respectively. The loss-function is discussed in a number of papers in the literature on predicting financial crisis, see for example Demirguc-Kunt and Detragiache (1999) and Bussiere and Fratzscher (2002).

Table 5.2.a: Competing-risks model: France

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
Event (E1 = financial distress)	Correct call of event: 75 pct. (1,280 out of 1,703)	Type 1 error: Missing prediction: 25 pct. (423 out of 1,703)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)	Type 2 error: Wrong signal: 27 pct. (29,347 out of 106,830)	Correct call of non-event: 72 pct. (77,213 out of 106,830)

[27] This means that in the 1) French case, the cut-off level is 0.01594 (=1703/106830), in the 2) Spanish case, the cut-off level is 0.00238 (=180/75686) and in the 3) Italian case, the cut-off level is 0.00159 (=155/97577).

Table 5.2.b: Competing-risks model: Italy

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
Event (E1 = financial distress)	Correct call of event: 88 pct. (137 out of 155)	Type 1 error: Missing prediction: 12 pct. (18 out of 155)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)	Type 2 error: Wrong signal: 29 pct. (28,283 out of 97,577)	Correct call of non-event 71 pct. (69,294 out of 97,577)

Table 5.2.c: Competing-risks model: Spain

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
Event (E1 = financial distress)	Correct call of event: 76 pct. (137 out of 180)	Type 1 error: Missing prediction: 24 pct. (43 out of 180)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)	Type 2 error: Wrong signal: 32 pct. (24,114 out of 75,686)	Correct call of non-event: 68 pct. (51,572 out of 75,686)

6. Results: The Country Models compared with a Pooled Model

The data from the three country models is pooled, and a pooled country model is estimated. The significant variables and their sign, and its predictive ability is compared to the three country models in order to assess the differences in the determinants of financial distress and in the predictive ability of the two model set ups. The results from the specification of the model as a pooled model is interesting in the light of Basel II, which allows for banks to pool their data with other banks in order to overcome their data shortcomings (as valid estimates of the probability of default for individual banks require a considerable amount of data). A number of international data pooling projects have emerged, where banks from various countries pool their data. Because of this development and as, furthermore, many credit institutions in Europe have cross-border activities, the choice between setting up individual country credit-scoring models or a common credit-scoring model is relevant when calculating capital requirements for banks. Therefore, the determinants of corporate failure in French, Italian and Spanish firms are investigated by the estimation of individual country credit-scoring models and a common credit-scoring model. Borup, Kurek and Rommer (2005) take the analysis further and calculate the capital requirements in each country and for the three countries as a whole (i.e. interpreted as one

banking portfolio) using different credit-scoring models. The implications for the resulting capital requirements for the portfolio which only entails loans in each country individually and for the three countries as whole (i.e. interpreted as one banking portfolio) are then discussed in their paper. In this paper the focus is not on the resulting capital requirements, but on the factors that drive financial distress. These are discussed using two extreme set ups, namely individual country credit-scoring models and a common credit-scoring model. The common credit-scoring model is estimated using the same explanatory variables as in the country credit-scoring models. It is not extended with either dummies for each country, or with interactions between dummies for each country and some of the already included explanatory variables (e.g. profitability). It has been chosen to show the two extremes, namely the country credit-scoring model and a common credit-scoring model.

6.1 Parameter Estimates

The significance and sign of the parameter estimates in the country models are compared to the significance and sign of the parameter estimates in the pooled model. Table 6.1 shows that all core and proxy variables are significant in the pooled model, except the core variable loans to total assets. This result differs from the results in the country models.

Table 6.1: E1 hazard

	France	Spain	Italy	Pooled model
Age	0.00302*	0.00784	-0.0298*	0.00437*
Size	-0.3265*	0.0294	0.2978*	-0.3847*
Earnings ratio	-1.6054*	-1.1842*	-2.4314*	-1.5114*
Solvency ratio	-1.5832*	-0.9511*	-2.3606*	-1.5092*
Loans to total assets	0.3169*	1.4822*	-0.2008	0.1227
Legal form	0.0624	0.1166	0.4787*	-0.8705*
Shareholders	-0.0214	0.0271	0.0258	-0.0775*
Subsidiaries	0.0355	-0.0516	-0.0954	-0.0477*
Bvd_indep_a	0.1973	0.1721	-0.0391	0.8610*
Bvd_indep_b	-0.0924	0.0223	-0.2694	0.3020*
Bvd_indep_c	-0.3653*	-0.3174	0.0127	0.6529*

Note: For details on the country models the reader is referred to table 5.1.a. The pooled country model is estimated using the same variables as in the country models. Again, a significance level of 5 pct. is chosen.

First the core variables are compared in the two model set ups. The country model that resembles the pooled model the most (in terms of what predictors of financial distress are significant and their sign) is the French country model. Except of the loans to total assets ratio, which is not significant in the pooled model, but is significant in the French country model, all core variables are significant in both the French country model and the pooled model and they have the same sign and are of similar size in the two set ups. When the pooled model is compared to the Italian and the Spanish country models, only the earnings ratio and the solvency ratio are significant and have the same sign in both model set ups.

All proxy variables are significant in the pooled model. Only one proxy variable was significant in the French case, none of the proxy variables were significant in the Spanish case and only one proxy variable was significant in the Italian case.

The proxy variables that are significant in the pooled model, but are insignificant in all the country models, are the ownership variables bvd_indep_a and bvd_indep_b and the variables number of shareholders that a firm has registered and number of subsidiaries a firm has registered. Bvd_indep_a and bvd_indep_b are significant and have a positive sign. This indicates that firms whose ownership is not so concentrated are more likely to enter financial distress. The variables number of shareholders that a firm has registered and number of subsidiaries a firm has registered have a negative sign, indicating that a larger number of registered shareholders and subsidiaries lowers the likelihood of financial distress. No hypotheses were set up on the effects of these variables in section 3.2.

The sign of two of the significant proxy variables in the pooled model seem puzzling. This is the sign on the ownership variable bvd_indep_c and the sign on the legal status variable. These variables have different signs compared to the two country models, where they were also significant.

The proxy variable bvd_indep_c is significant and has a negative sign in the French credit-scoring model, indicating that concentration of ownership leads to a lower likelihood of entering financial distress. This result confirms that it is desirable to concentrate ownership among few individual, c.f. Bennedsen and Wolfenzon (2000). It is therefore puzzling that the ownership variable bvd_indep_c is significant and has a positive sign in the pooled model.

The proxy variable legal form is significant and has a negative sign in the Italian credit-scoring model, indicating that private limited liability companies have a larger probability of moving into financial distress. The variable was insignificant for France and Spain. In the pooled model the variable is significant and has a negative sign, indicating that private limited liability companies are less likely to enter financial distress. This is puzzling. One would think that private limited liability companies would be more risky, as they would have less share capital to loose compared with public limited liability companies.

The overall result is that the pooled model delivers parameter estimates, which in terms of significance and sign differ to quite an extent from all the country credit-scoring models. The implication of this is

that country credit-scoring should be estimated. It does not make sense to pool the data and estimate a common credit-scoring model.

6.2 Discriminatory Power

Table 6.2.a shows that, overall, the discriminatory power of the pooled country model is very alike the individual country models. The pooled country model correctly classifies 74 pct. of the financially distressed firms as financially distressed, and 72 pct. of the active firms as active.[28] The overall result of table 6.2.a hides important differences between the two model set ups. Table 6.2.b shows that only 66 pct. (=(34,187+152,218)/282,131)) of the predictions are the same in the two model set ups. In quite a number of cases the country models predict an event, when the pooled country model does not predict an event (49,135) and vice versa.

Table 6.2.a: Competing-risks model for France, Italy and Spain (pooled country model)

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
Event (E1 = financial distress)	Correct call of event: 74 pct. (1,498 out of 2,038)	Type 1 error: Missing prediction: 26 pct. (540 out of 2,038)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)	Type 2 error: Wrong signal: 28 pct. (79,280 out of 280,093)	Correct call of non-event: 72 pct. (198,131 out of 280,093)

Table 6.2.b: A comparison of the predictions of the three country models and the pooled country model

		Pooled country model		Sum
		Model prediction: Event	Model prediction: Non-event	
Country models	Model prediction: Event	34,187	49,135	83,322
	Model prediction: Non-event	46,591	152,218	198,809
Sum		80,778	201,353	282,131

[28] The cut-off level is 0.00728 (=2038/280093).

Table 6.2.c: Model predictions and actual events

Actual event		Model prediction	
		Pooled model = No-event Country models = Event	Pooled model = Event Country models = No-event
Event	E1	245	186
No-event	E2	174	90
	E3	326	178
	E4	108	120
	Active	48,282	46,017
Sum		49,135	46,591

Table 6.2.d: Model predictions of the E1 event split up on country

		Model predictions			
		Pooled model = No-event Country models = Event		Pooled model = Event Country models = No-event	
		Number of observations	Number of observations as a percentage of the number of E1 firms in the respective countries	Number of observations	Number of observations as a percentage of the number of E1 firms in the respective countries
E1	Spain	74	41 pct.	15	8 pct.
	France	98	8 pct.	165	13 pct.
	Italy	73	47 pct.	6	4 pct.
	Sum	245	15 pct.	186	12 pct.

Details on the cases, where the models have different predictions, are provided in table 6.2.c, which tabulates the predictions and the actual events. The first column of table 6.2.c splits up the 49,135 cases (where the pooled model predicts a no-event and the country models predict an event) up on the actual outcome, which is observed in the dataset. Where the pooled country model predicts a "no-event", the three country models predict financial distress in 245 firms that turn out to be financially distressed. The other way around it is 186 cases. In 14 pct. (59 out of 431) of the cases, where the models have different predictions, the country models do a better job. The three country models are better at predicting Spanish and Italian firms in financial distress, whereas the pooled country model predicts a larger amount of the French firms correctly. This can be seen from 6.2.d, which shows a detailed breakdown of the first row in table 6.2.c, or, in other words, the E1 predictions (when they are not the same in the two set ups) split up on country. According to the table, 41 pct. of the Spanish firms that end up in financial distress are predicted to end up in financial distress, when the country models are used. Had one used the pooled country model, one would not have predicted these firms to end up in

financial distress. Instead 15 other Spanish firms (8 pct. of all Spanish firms in financial distress), which were not predicted to end up in financial distress when the country models were used, would be predicted to end up in financial distress, when the pooled model is used. When measured on the number of firms that are predicted to enter financial distress and actually do enter financially distress, the three country models do a better job. The difference between the two set ups in terms of financial distress prediction of Spanish firms is 59 firms, corresponding to 33 pct. of all the Spanish firms that end up in financial distress. The same picture concerns the Italian case, whereas, in the French case, the result is the opposite. Here the pooled country model does a better job.

The finding, that the pooled country model scores better than the individual country model for France, but worse for Spain and Italy, may not be particularly surprising, given how the cut-off levels are chosen. Indeed, by comparing the individual cut-off levels (0.016 for France, 0.0024 for Spain and 0.0016 for Italy) with the pooled country cut-off level (0.0073), one can already "guess" that the results will improve for France (the threshold is set "too low") while they will worsen for the other two countries (the threshold is set "too high"). Instead of choosing the cut-off level as number of firms in financial distress to all other firms, one could have chosen to match the probability that type I or type II errors occur (or to have used other methods), but this is not done here. The way the cut-off level is chosen should depend on the loss-function of the "agent" that is using the model for predicting. No matter what cut-off is used, the results will always be conditional on the chosen method.

7. Conclusion

This paper investigated the determinants of corporate failure in Italian, Spanish and French SMEs using a dataset from Bureau van Dijk. The great virtue of the data set is that it enables us to make cross-country comparisons. On the negative side it should be mentioned that the data set, when looking at each country individually, is not as good as some of the data sets used in the individual country studies (in the sense that a number of firms drop out of the panel with no explanation).

Italy, Spain and France are countries, which in important aspects are fairly alike. They all belong to Continental Europe, they are all members of the European Monetary Union and they are inspired by the same legal tradition. Furthermore, despite the deregulation and liberalisation process of the financial systems, which took place in the countries in the 1980s and 1990s, banks are still very important sources of financing. Based on this, a priori, one may have thought that the same factors were likely to drive financial distress in the three countries.

The estimations of country credit-scoring models show that there are some similarities across countries, but also that there are important differences in the determinants of financial distress. The core variables that behave similarly across countries are the earnings ratio and the solvency ratio. They are significant and have a negative sign in all countries. The proxy variables that behave similarly across countries are the number of subsidiaries a firm has registered, the number of shareholders a firm has registered, the ownership variables bvd_indep_a and bvd_indep_b. These variables are insignificant in all countries.

The variables, whose effect differs between the countries in terms of whether or not they are significant or what sign they have, are the loans to total assets ratio, size, age, legal form and the ownership variable bvd_indep_c.

The data from the three country models is pooled, and a pooled country model is estimated. As valid estimates of the probability of default using the Basel II framework require a considerable amount of data, Basel II allows for banks to pool their data with other banks in order to overcome their data shortcomings, and so a number of international data pooling projects have emerged, where banks from various countries pool their data. Because of this development and as, furthermore, many credit institutions in Europe have cross-border activities, the choice between setting up individual country credit-scoring models or a common credit-scoring model is relevant, when calculating capital requirements for banks. In order to shed light on these issues, the determinants of corporate failure in French, Italian and Spanish firms are investigated by the estimation of individual country credit-scoring models and a common credit-scoring model. The significant variables and their sign, and the predictive ability of the common credit-scoring model is compared to the three country models in order to assess the differences in the determinants of financial distress and in the predictive ability of the two model set ups.

The result from the individual credit-scoring models are confirmed by the estimation of the pooled model, which shows that the pooled model delivers parameter estimates that differ to quite an extent from all the country credit-scoring models. The comparison of the core variables in the pooled model with the individual country models show that the country that resembles the pooled model the most (in terms of what predictors of financial distress are significant and their sign) is France. The comparison of the proxy variables in the pooled model with the individual country models showed that there were no similarities between the country models and the pooled model. The overall conclusion is that the pooled model delivers parameter estimates that differ to quite an extent from the all the country credit-scoring models. A comparison of the predictive ability of the pooled model and the country models hides important differences between the two model set ups, namely that the pooled model does better for France than the French credit-scoring model, but worse for Spain and Italy than the Spanish and the Italian credit-scoring models. It can seem a bit surprising that the French country model performs worse for France than the pooled country model, but it is not so surprising, when one takes a closer look on how the cut-off levels are chosen. It is therefore important to highlight that the result is conditional on the chosen cut-off levels. It was chosen to use the proportion of firms in financial distress to all other firms.

The overall conclusion from the analysis is that not even in this case, where the analysed sample of countries is fairly homogenous, does it make sense to estimate one common credit-scoring model. This is an important conclusion, which has implications for at least two policy areas, namely financial stability analysis and Basel II.

8. LITERATURE

Alderman, H., J. R. Behrman, H-P. Kohler, K. A. Maluccio and S. C. Watkins, 2000. *Attrition in Longitudinal Household Survey Data. Some Tests for Three Developing-Country Samples.* Policy Research Working Paper no. 2447, The World Bank, September 2000.

Allison, P. D., 1982. Discrete-time Methods for the Analysis of Event Histories. In Samuel Leinhardt (ed.), *Sociological Methodology* 1982, San Francisco: Jossey-Bass, pp. 61-98

Allison, P. D., 2001. *Logistic Regression using the SAS System. Theory and Application*. Cary, NC, USA: SAS Institute

Altman, E. I., 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, vol. 23, no. 4, pp. 589-609

Altman, E. I., G. Marco and F. Varetto, 1994. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, vol. 18., pp. 505-529

Ayadi, R., 2004. *Impact of the New Basel Capital Accord on SME Financing.* In The New Basel Capital Accord and the Future of the European Financial System, Report of a CEPS Task Force no. 51, Rapporteurs: R. Ayadi and A. Resti, April 2004

Bardos, M., 1998. Detecting the risk of company failure at the Banque de France. *Journal of Banking & Finance*, vol. 22, pp. 1405-1419

Bardos, M., 2001. *Recent developments in the Banque de France's scoring method.* Banque de France Bulletin Digest, no. 93, September 2001

Basel Committee on Banking Supervision, 2004. *International Convergence of Capital Measurement and Capital Standards. A Revised Framework.* Bank for International Settlements, June 2004

Beaver, W., 1966. Financial Ratios as Predictors of Bankruptcy. *Journal of Accounting Research*, vol. 6, pp. 71-102

Bennedsen, M., 2004. *The Governance of Closely Held Corporations: Ownership Structure, Family Management and the Role of the Corporate Board.* Manuscript presented at the Annual Meeting in Nationaloekonomisk Forening, 2004.

Bennedsen, M. and D. Wolfenzon, 2000. The balance of power in closely held corporations. *Journal of Financial Economics*, vol. 58, pp. 113-139

Bhattacharjee, A., C. Higson, S. Holly and P. Kattuman, 2004. *Business Failure in UK and US Quoted Firms: Impact of Macroeconomic Instability and the Role of Legal Institutions.* Unpublished manuscript, January 2004.

Borup, L., Kurek, D. and A. D. Rommer, 2005. *Capital requirements for corporate default risk: The consequences of pooling data from French, Italian and Spanish SMEs*. Mimeo

Bunn, P., and V. Redwood, 2003. *Company accounts based modelling of business failures and the implications for financial stability*. Bank of England Working Paper No. 210

Bureau van Dijk, 2004. *Ownership Database.* March 2004

Bussiere, M. and M. Fratzscher, 2002. *Towards a new early warning system of financial crises.* European Central Bank working paper series. Working paper no. 145, May 2002.

Cappiello, S. and G. Marano, 2003. The Reform of the Legal Framework for Italian Enterprises and the 2003 Company Law. *International Company and Commercial Law Review*, n. 6, 2003

Cifarelle, D. M. and F. Corielli, 1988. Business Failure Analysis. A Baysian Approach With Italian Firm Data. *Studies in Banking and Finance,* no. 7, pp. 73 - 89

Corcóstegui, C., L. González-Mosquera, A. Marcelo and C. Trucharte, 2003. *Analysis of procyclical effects on capital requirements derived from a rating system.* Paper presented at Basel Committee/Banca d'Italia workshop, 20-21 March 2003.

Creditreform, 2002. *Insolvencies in Europe 2001/2*. A Survey by the Creditreform Economic Research Unit

Creditreform, 2003. *Insolvencies in Europe 2002/3.* A Survey by the Creditreform Economic Research Unit

Danmarks Statistik, 2002. *Rapport fra projektgruppen om Demografi.* Danmarks Statistik, Erhvervsstruktur, December 2002.

Demirguc-Kunt, A. and E. Detragiache, 1999. *Monitoring Banking Sector Fragility: A Multivariate Logit Approach.* International Monetary Fund Working Paper, WP/99/

Davydenko, S. A. and J. R. Franks, 2004. *Do bankruptcy codes matter? A study of defaults in France, Germany, and the UK.* Unpublished manuscript, June 2004.

Dietsch, M. and J. Petey, 2002. The credit risk SME loans portfolios: Modeling issues, pricing, and capital requirements. *Journal of Banking & Finance*, vol. 26, pp. 303-322

Diggle, P. and M. G. Kenward, 2004. Informative Drop-out in Longitudinal Data Analysis. *Applied Statistics*, vol. 43, no. 1, pp. 49-93

Dyrberg, A., 2004. *Firms in Financial Distress: An Exploratory Analysis.* WP no. 17, Danmarks Nationalbank

European Central Bank, 2002. *Report on Financial Structures.* European Central Bank

European Central Bank, 2004. *Financial Stability Review, December 2004.* European Central Bank

European Commission, 2003a. *Bankruptcy and a fresh start: stigma on failure and legal consequences of bankruptcy.* November 2003

European Commission, 2003b. *Bankruptcy and a fresh start: stigma on failure and legal consequences of bankruptcy. Italian Country Study.* November 2003

European Commission, 2003c. *Bankruptcy and a fresh start: stigma on failure and legal consequences of bankruptcy. Spanish Country Study.* November 2003

European Commission, 2003d. *Bankruptcy and a fresh start: stigma on failure and legal consequences of bankruptcy. French Country Study.* November 2003

European Commission, 2003e. *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises.* Official Journal of the European Union, L 124/36, 2003

European Committee of Central Balance Sheet Offices (2000), *Corporate Finance in Europe from 1986 to 1996.* Own Funds Working Group, European Committee of Central Balance Sheet Offices, 2000.

Eurostat, 2004. *Business Demography. Recommendations Manual.* Eurostat, 2004

Feelders, A. J., 2003. *An Overview of Model Based Reject Inference for Credit Scoring.* Unpublished manuscript

Fernandez, A. I, 1988. A Spanish Model for Credit Risk Classification. *Studies in Banking and Finance*, vol. 7, pp. 115-125

Fitzgerald, J., P. Gottschalk and R. Moffitt, 1998. An Analysis of Sample Attrition in Panel Data. *The Journal of Human Resources*, vol. 33, no. 2, pp. 251-299

Fitzmaurice, G. M., A. F. Heath and P. Clifford, 1996. Logistic Regression Models for Binary Panel Data with Attrition. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 159, no. 2, pp. 249-263

Greene, W., 2003. *Econometric Analysis*. New Jersey, USA: Prentice Hall

Harhoff, D., K. Stahl and M. Woywode, 1998. Legal Form, Growth and Exit of West German Firms: Empirical Results for Manufacturing, Trade and Service Industries. *Journal of Industrial Economics*, vol. 46, pp. 453-488

Hayden, E., 2003. *Are Credit Scoring Models Sensitive With Respect to Default Definitions? Evidence from the Austrian Market.* Unpublished manuscript, 2003

Henley, A., 1998. Residential mobility, housing equity and the labour market. *The Economic Journal*, 108, pp. 414-427

Hermalin, B. E. and M. S. Weisbach, 2003. Boards of Directors as an Endogenously Determined Institution: A Survey of the Economic Literature. *FRBNY Economic Policy Review*, April 2003.

Hunter, J. and N. Isachenkova, 2000. *Failure Risk: a Comparative Study of UK and Russian Firms.* Department of Economics and Finance Brunel University, Discussion Paper 00-1

Jenkins, S., 1995. Easy Estimation Methods for Discrete-Time Duration Models. *Oxford Bulletin of Economics and Statistics*, vol. 57, no. 1, pp. 129-138

Jenkins, S., 2003. Survival Analysis. Unpublished manuscript

Jiménez, G. and J. Saurina, 2004. *Collateral, type of lender and relationship banking as determinants of credit risk.* WP no. 0414, Banco de España

Köke, J., 2001. *Determinants of acquisition and failure: Stylized facts and lessons for empirical studies.* Discussion Paper No. 01-30, ZEW

Lando, D., 2004. *Credit Risk Modeling. Theory and Applications.* Princeton Series in Finance

La Porta, R., F. Lopez-De-Silanes, A. Shleifer and R. Vishny, 1998. Law and Finance. *The Journal of Political Economy*, vol. 106, no. 6, pp. 113-1155

Memorandum of Understanding on the exchange of information among national credit registers for the purpose of passing it on to reporting institutions, 20 February 2003

Moody's Investors Service, 2000. *Moody's RiskCalc For Private Companies: Moody's Default Model. Rating Methodology*, May 2000

Moody's Investors Service, 2001a. *Moody's RiskCalc For Private Companies: Spain. Rating Methodology*, July 2001

Moody's Investors Service 2001b. *Moody's RiskCalc For Private Companies: France. Rating Methodology*, December 2001

Moody's Investors Service 2002. *Moody's RiskCalc For Private Companies: Italy. Rating Methodology*, October 2002

Ohlson, J. A., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, vol. 19, pp. 109-131

Ooghe, H. and S. Balcaen, 2002. *Are failure prediction models transferable from one country to another? An empirical study using Belgian financial statements.* Vlerick Working Papers 2002/5

Prantl, S., 2003. *Bankruptcy and Voluntary Liquidation: Evidence for New Firms in East and West Germany after Unification.* Discussion Paper No. 03-72, ZEW

Rommer, A. D. *Testing the Assumptions in Credit-scoring Models.* Mimeo

Scheule, H., 2003. *Prognose von Kreditausfallrisiken.* Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenshcaften eingereicht an der Wirtschaftswissenschartlichen Fakultät der Universität Regensburg, vorgelegt von Harald Scheule, Schongau, April 2003

Shumway, T., 2001. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, vol. 74, no. 1, pp. 101-124

Wooldridge, J. M., 2002. *Econometric Analysis of Cross section and Panel data*. Cambridge (Mass.), USA: The MIT Press

Wooldridge, J. M., 2003. *Introductory Econometrics. A Modern Approach.* Ohio, USA: Thomson South-Western

9. Appendix: Other Studies

Table 9.a: Studies using French data

Study	Data	Method	Default definition
Bardos (2001) (see also Bardos (1998)) FR	The Banque de France FIBEN database (The database is almost exhaustive for firms with a turnover exceeding FRF 5 million. Some smaller firms are recorded in the database, particularly if their debt exceeds FRF 2.5 million. The debt of the firms in the database is proportionally higher than that seen in the exhaustive population surveyed by INSEE (French Institute of Statistics and Economics Studies)). The study uses representative data on French companies in the FIBEN database, which have a turnover exceeding EUR 0.75 million or whose bank loans exceeded (five times) a risk declaration threshold. Period: 1991 – 2000	Probability of default (PD) estimation: Score (discriminant analysis). A probability of failure is associated with each value of the score	Failure = opening of legal proceedings. It is noted that the opening of legal proceedings mainly concerns small- and medium-sized companies rather than large companies, as large companies, which are experiencing difficulties, rarely submit their balance sheet to the registrars of Commercial Courts (they generally restructure and/or negotiate their debts)
Dietsch and Petey (2002) FR	The analysis is done on 224.000 French SMEs. Data is provided by Coface SCRL. Period: 1995 –1999	PD estimation: Ordered probit model and a model using a gamma distribution.	Default corresponds to bankruptcy
Moody's Investors Service (2001b) FR	1,323,754 financial statements and 25,229 defaults from 253,268 French private companies. Only firms with a turnover of at least 0.5 million euro. Period: 1990 – 1999	PD estimation: Binomial logit model	Bankruptcy or insolvency
This paper ES FR IT	Italy: 97,733 SME firm-years Spain: 75,866 SME firm-years France: 108,533 SME firm-years Data is provided by Bureau van Dijk (Amadeus database). Period covered for all countries: 2000 – 2002	PD estimation: Competing-risks model	Financially distressed firms include firms that are bankrupt, active (receivership) and active (default of payments)

Table 9.b: Studies using Spanish data

Study	Data	Method	Default definition
Jiménez and Saurina (2004) ES	Credit Register of the Bank of Spain (CIR). Monthly information on all loans granted by credit institutions in Spain for a value of 6,000 euros. The data has been subjected to various filters, e.g. only loans to companies above a threshold of 24,000 euros are used. Data on over 3 million loans are analysed. Data from December. Five years are included: 1987, 1990, 1993, 1997, 2000	Probability of default (PD) estimation: Binomial logit model	Default on payment is considered to have occurred when, three months after the date of maturity, the debt balance remains unpaid or when there are reasonable doubts as to its repayment
Corcóstegui, González-Mosquera, Marcelo and Trucharte (2003) ES	Credit Register of the Bank of Spain (CIR). Bank of Spain's Central Financial Database (CBBE). Private database SABE. The raw data consists of the obligors that are found in all three databases. Various filters are used, including a minimum size threshold in terms of annual volume of sales equal to 9 million euro. 73,321 obligors are analysed. Period: 1993 – 2000	PD estimation: Binomial logit model	Default on payment is considered to have occurred when, three months after the date of maturity, the debt balance remains unpaid or when there are reasonable doubts as to its repayment
Fernandez (1988) ES	A matched sample of 70 firms. Data set was drawn from the files of loan clients of a Spanish bank.	PD estimation: Univariate analysis, factor analysis by principal components, discriminant analysis	Failure is defined as credit applicants that the bank classified as insolvent (e.g. operations whose ability to repay a loan is doubtful, and thus, not only firms which have failed in a legal sense (bankrupt firms))
Moody's Investors Service (2001a) ES	569,181 financial statements and 2,265 defaults from 140,790 companies. Only firms with a turnover of at least 0.5 million euro. The database is developed by Equifax. Period: 1992 – 1999	PD estimation: Binomial logit model	Bankruptcy or insolvency
This paper ES FR IT	Italy: 97,733 SME firm-years Spain: 75,866 SME firm-years France: 108,533 SME firm-years Data is provided by Bureau van Dijk (Amadeus database). Period covered for all countries: 2000 – 2002	PD estimation: Competing-risks model	Financially distressed firms include firms that are bankrupt, active (receivership) and active (default of payments)

Table 9.c: Studies using Italian data

Study	Data	Method	Default definition
Cifarelli and Corielli (1988) IT	A sample of 27 unsound and 196 sound firms is used (reference is taking in a 'major' Italian bank)	Probability of default (PD) estimation: A bayesian variant to discriminant analysis	Default corresponds to bankruptcy
Altman, Marco and Varetto (1994) IT	A matched sample of all together 1,000 healthy and unsound industrial Italian firms. Focus is on medium and small sized businesses. For this reason, companies with sales of more than 100 billion lira (i.e. 60 million dollars) have been excluded from the sample. The sample is obtained from the Central dei Bilanci (CB), which is an organization established by the Banca d'Italia. Period: 1982 – 1992	PD estimation: Linear discriminant analysis, logit analysis and neural networks	Defaults correspond to 1) bankruptcy, 2) firms that are wound up in temporary receivership, 3) firms that had stated they were in dire straits with regard to their payments to the banks.
Moody's Investors Service (2002) IT	124,937 financial statements and 958 defaults from over 52,329 Italian private companies. Only firms with a turnover of at least 0.5 million euro. Period: 1995 – 1999	PD estimation: Binomial logit model	Defaults are based on three sources of information 1) Sponsor bank experience, 2) provisioning data and 3) insolvency statistics
This paper ES FR IT	Italy: 97,733 SME firm-years Spain: 75,866 SME firm-years France: 108,533 SME firm-years Data is provided by Bureau van Dijk (Amadeus database). Period covered for all countries: 2000 – 2002	PD estimation: Competing-risks model	Financially distressed firms include firms that are bankrupt, active (receivership) and active (default of payments)

10. Appendix: Legal status codes

Table 10.a: French legal status codes

Situation normale, avec comptes	Active
Pas d'obligation de dépôt (NC)	Active
Plan de continuation (PC)	Active (receivership)
Redressement judiciaire par jugement (RJ)	Active (receivership)
Liquidation par jugement (LJ)	Bankruptcy
Absorption, fusion ou rachat (AF)	Dissolved (merger)
Liquidation à l'amiable (LA)	In liquidation
Cessation d'activité (CE)	Inactive (no precision)
Cessation de paiement (CP)	Active (default of payments)

Source: Bureau van Dijk

Table 10.b: Italian legal status codes

Ditta attiva	Active
Ditta in liquidazione	In liquidation
Ditta in fallimento	Bankruptcy
Ditta sospesa	Active (dormant)
Ditta inattiva	Inactive (no precision)
Ditta cassata	Dissolved
Ditta cessata per trasferimento	Dissolved (merger)

Source: Bureau van Dijk

Table 10.c: Spanish legal status codes

Activa	Active
Suspension de pagos	Active (default of payments)
Quiebra	Bankruptcy
Disuelta	Dissolved
Absorbida	Dissolved (merger)
Extinguida	Dissolved
Inactiva	Inactive (no precision)

Source: Bureau van Dijk

11. Appendix: Sample Selection in Details

The raw data is split up according to 6 different codes: C1 = Consolidated statement with no unconsolidated companion, C2 = Consolidated statement with an unconsolidated companion, U1 = Unconsolidated statement with no consolidated companion, U2 = Unconsolidated statement with a consolidated companion, LF = Limited financial data, probably unconsolidated, N.A. = No financial data available. In order to avoid double accounting, only companies with the codes U1 and U2 are analysed (For Italy, Spain and France there are no financial statements with the LF code). In the raw data there are 3,777,275 firm-year observations covering the period 2000 – 2002 (FR: 2,135,745, IT: 461,985, ES: 1,179,545).

After the exclusion of financial institutions and non-financial holding companies, there are 3,629,353 firm-year observations left (FR: 2,019,361, IT: 459,611, ES: 1,150,381).

A panel dataset is constructed. If a company hands in two financial statements, only the last financial statement is included in the estimations (FR: 2,019,216, IT: 459,577, ES: 1,154,378). Active companies are excluded if they hand in a financial statement in 2000 and 2002 (and not in 2001) (FR: 1,992,780, IT: 442,608, ES: 1,153,092).

The first year a specific company enters the dataset, it is ensured that it has at least 10 employees and total assets equal to or above 2 million euro. Remaining are 383,845 firm-year observations (FR: 137,183, IT: 146,360, ES: 100,302). Companies with 250 employees or more and companies with a total balance sheet of more than 43 million euro are excluded, making the database a total of 345,078 firm-year observations (FR: 120,758, IT: 135,010, ES: 89,310).

Finally, various corrections are made to the database (e.g. firms with illogical variables, such as short-term debt less than zero and a solvency ratio larger than 100 pct., are excluded) (FR: 118,395, IT: 134,038, ES: 81,162).

Only public limited liabilities and private limited liabilities are analysed. (FR: 113,421, IT: 126,056, ES: 80,175).

Firms with missing variables on any of the explanatory variables are excluded. Remaining are 299,746 firm-year observations (FR: 113,420, IT[29] [30]: 106,166, ES: 80,160).

The attritioners are excluded making the database a total of 282,131 firm-year observations (FR: 108,533, IT: 97,732, ES: 75,866).

[29] Most of the Italian firms are excluded because there is no information on the incorporation date (and so the age of the company cannot be constructed).

[30] BvD has provided me with further information in the Italian firms. As of 2002, their Italian data provider includes information on firms with a turnover between 500.000 and 1.000.000 euro (these firms were not included before). In the dataset 927 of the Italian firms are new firms in 2002 with a turnover less than 1.000.000 euro.

12. Appendix: Detailed analysis of the Attritioners

12.1 Introduction

The attritioners are analysed in detail in this appendix. The means of the attritioners and the non-attritioners are compared. To investigate whether the means of various variables are significantly different, t-tests are performed and results are reported. The null hypothesis is that there is no difference between the means. The alternative hypothesis is that they are different. In order for the t-test to be valid, it is necessary that the responses are independent of each other, and the observations should come from a normal distribution. If the sample sizes are equal, the procedure is fairly robust to violations of the normality assumption. Here the sample sizes are not equal, and so the nonparametric Wilcoxon's test is also performed. The result of this test is very similar to the t-test (results are not reported). The t-test assumes equal variances. A test that allows for non-equal variances is the Satterthwaite test. This test is also performed and results are reported. Thereafter attrition probits are estimated. The variables that drive attrition are discussed. Finally, a robustness check is performed.

12.2 Testing for Equal Means of the Explanatory Variables

France: 2000 sample

The E5 firms and non-E5 firms resemble each other (table 12.2.a). Out of the 22 core, proxy and control variables, which are examined, only 3 – 4 core and proxy variables (depending on the test) and 3 – 4 control variables (depending on the test), are rejected to have equal means using both the t-test and the Satterthwaite test for the 2000 sample. The core variables that are rejected to have equal means are the solvency ratio, age and size (only t-test), and the proxy variable, which is rejected to have equal means, is the dummy legalform. The 4 sector affiliation dummies that are rejected to have equal means are: dumman, dumcon, dummin (only Satterthwaite test) and dumpub.

It is worth noticing that the solvency ratio is larger for the attritioners than for the non-attritioners. The solvency ratio is hypothesized to affect the firm that enter financial distress negatively, and so, based on this variable, there is no sign of the E5 firms being weaker than the non-E5 firms. The age of the E5 firms is lower than the age of the non-E5 firms and the E5 firms are larger than the non-E5 firms. Both age and size were hypothesized to affect firms in financial distress negatively. Private limited liability companies are more likely to be among the E5 firms compared to the non-E5 firms. This variable, which is a proxy for the willingness to take on risk, was hypothesized to affect the firms in financial distress positively.

The variables, which cannot be rejected to have equal means, are the earnings ratio, the loans to assets ratio, the size of the company (only Satterthwaite test), the number of shareholders, the number of subsidiaries, bvd_indep_a, bvd_indep_b, bvd_indep_c, IT dummy, dumorg, dumfar, dummin, dumene, dumbus, dumpub, dumtrahot and dumtra. Some of these variables were hypothesized to affect the firms that enter financial distress in a certain direction. These were the earnings ratio, the loans to assets ratio, size and the bvd_indep_c. For other variables (the number of shareholders, the number of subsidiaries, bvd_indep_a and bvd_indep_b) no particular hypothesis was set up. It was left to the estimations to show whether or not there was an effect.

Table 12.2.a: France: Descriptive Statistics and tests: 2000 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard deviation	Mean	T-test	Satterthwaite
Age	20.05155	16.41487	22.44933	18.26483	-2.39778	0.0001	<0.0001
Size	8.574916	0.722556	8.52668	0.715432	0.048236	0.0496	0.0520
Earnings ratio	0.10225	0.157116	0.103828	0.135496	-0.00158	0.7356	0.7691
Solvency ratio	0.317633	0.271554	0.298217	0.245842	0.019417	0.0218	0.0370
Loans to assets ratio	0.077069	0.125895	0.074109	0.112584	0.00296	0.4453	0.4923
Legal form	0.159221	0.366092	0.11863	0.323358	0.040591	0.0003	0.0012
Shareholders	1.280641	1.159446	1.354388	1.26352	-0.07375	0.0884	0.0649
Subsidiaries	0.483391	1.059042	0.489625	1.334851	-0.00623	0.8912	0.8649
Bvd_indep_a	0.033219	0.17931	0.029331	0.168736	0.003888	0.5028	0.5271
Bvd_indep_b	0.027491	0.163604	0.031321	0.174188	-0.00383	0.5211	0.4962
Bvd_indep_c	0.681558	0.466139	0.677749	0.467346	0.003809	0.8123	0.8119
It dummy	0.040092	0.196287	0.030538	0.172066	0.009553	0.1072	0.1552
Dumman	0.367698	0.482455	0.326003	0.468756	0.041694	0.0096	0.0119
Dumorg	0.019473	0.13826	0.02075	0.14255	-0.00128	0.7939	0.7880
Dumcon	0.050401	0.218896	0.080261	0.271701	-0.02986	0.0013	<0.0001
Dumfar	0.010309	0.101068	0.01155	0.106849	-0.00124	0.7348	0.7212
Dummin	0.003436	0.058554	0.008189	0.090125	-0.00475	0.1214	0.0205
Dumene	0.001145	0.033845	0.002643	0.05134	-0.0015	0.3918	0.2057
Dumbus	0.145475	0.352782	0.129038	0.335247	0.016438	0.1538	0.1744
Dumpub	0.008018	0.089237	0.021631	0.145479	-0.01361	0.0060	<0.0001
Dumtrahot	0.33677	0.472876	0.342219	0.47446	-0.00545	0.7379	0.7372
Dumtra	0.057274	0.232498	0.057716	0.23321	-0.00044	0.9559	0.9558

Note: For more details on the variables the reader is referred to section 3.2. The following abbreviations are used: Dumfar = Farming, forestry and fishing, Dummin = Mining, Dumman= Manufacturing, Dumener = Energy, Dumcon = Construction, Dumtraho = Trade and hotel, Dumtra = Transport, Dumbus = Business service, Dumpub = Public service activities, Dumorg = Organisations. Subsidiaries: This variable measures the number of subsidiaries that a company has registered. Legal form: This dummy is equal to 1, if it is a private limited liability company, and equal to 0, if it is a public limited liability company. Shareholders: This variable measures the number of recorded shareholders. Independence indicators: Bvd_ind_a, bvd_ind_b and bvd_ind_c.

France: 2001 sample

Using the 2001 sample, out of the 22 core, proxy and control variables it is rejected that 3 core and proxy variables and 2 sector affiliation dummies have equal means, c.f. table 12.2.b. The core and proxy variables, which are rejected to have equal means, are the earnings ratio, the age of a company and the dummy variable legal form. The two sector affiliation dummies, which are rejected to have equal means, are dumcon and dumbus.

E5 firms have an earnings ratio that is smaller than the earnings ratio for the non-E5 firms in the 2001 sample. This could indicate that weak firms are more likely to be among the E5 firms. E5 firms are also younger than the non-E5 firms. As for the 2000 sample, in the 2001 sample private limited liability companies are more likely to be among the E5 firms compared to the non-E5 firms.

The variables, which cannot be rejected to have equal means, are the solvency ratio, the loans to assets ratio, the size of the company, the number of shareholders, the number of subsidiaries, bvd_indep_a, bvd_indep_b, bvd_indep_c, IT dummy, dumman, dumorg, dumfar, dummin, dumene, dumpub, dumtrahot and dumtra. Some of these variables were hypothesized to affect the firms that enter financial distress, e.g. the solvency ratio was hypothesized to have a positive sign and the loans to assets ratio was hypothesized to have a negative sign. No hypothesis was made on the effects of other variables, e.g. the number of shareholders and the number of subsidiaries.

France: Conclusion

In the French case, very few core and proxy variables are found to have different means. In the 2000 sample the mean of the solvency ratio, age, size (only t-test) and legal form was significantly different between E5 firms and non-E5 firms, and in the 2001 sample, the mean of the earnings ratio, age and legal form were significantly different between E5 firms and non-E5 firms. The solvency ratio is higher for E5 firms compared to non-E5 firms (2000 sample). The results in this section indicate that the E5 firms are stronger than the non-E5 firms. At the same time, the earnings ratio was found to be lower for the E5 firms than for the non-E5 firms, indicating that the E5 firms are weaker than the non-E5 firms (2001 sample). Age and size affect financial distress in opposite directions (2000 sample). For the 2001 sample, age of E5 and non-E5 firms cannot be rejected to have equal means with the E5 firms being younger than the non-E5 firms. Both age and size were hypothesized to affect the firms in financial distress negatively. Legal form is rejected to have equal means in both samples. In both samples private limited liability companies are more likely to be among the attritioners.

As most core and proxy variables cannot be rejected to have equal means for E5 and non-E5 firms, and as the effect of the solvency ratio (2000 sample) and the earnings ratio (2001 sample) give conflicting results concerning the potential bias of the estimates, the overall conclusion from the comparisons of means between the French E5 firms and non-E5 firms is that the means of the indicators, which are central to the study, do not differ in a systematically way between E5 firms and non-E5 firms.

Table 12.2.b: France: Descriptive Statistics tests: 2001 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard Deviation	Mean	T-test	Satterthwaite test
Age	20.4261	17.54955	22.40174	18.01933	-1.97564	<0.0001	<0.0001
Size	8.477076	0.722979	8.497622	0.734085	-0.02055	0.1763	0.1707
Earnings ratio	0.086599	0.19748	0.098183	0.154471	-0.01158	0.0004	0.0042
Solvency ratio	0.290889	0.29748	0.297292	0.265901	-0.0064	0.2489	0.2958
Loans to assets ratio	0.079453	0.124438	0.075101	0.118124	0.004352	0.0765	0.0905
Legal form	0.1751	0.380129	0.12983	0.336121	0.04527	<0.0001	<0.0001
Shareholders	1.318474	1.185226	1.337381	1.278889	-0.01891	0.4735	0.4438
Subsidiaries	0.465863	1.273963	0.448968	1.240744	0.016896	0.5117	0.5214
Bvd_indep_a	0.03494	0.183664	0.029311	0.168679	0.005629	0.1093	0.1372
Bvd_indep_b	0.032932	0.178494	0.031381	0.174348	0.001551	0.6682	0.6746
Bvd_indep_c	0.664659	0.472204	0.669298	0.470472	-0.00464	0.6342	0.6353
It dummy	0.038153	0.191603	0.033258	0.179312	0.004895	0.1896	0.2158
Dumman	0.314859	0.464553	0.31348	0.463914	0.00138	0.8859	0.8860
Dumorg	0.022892	0.149588	0.021611	0.145411	0.001281	0.6713	0.6789
Dumcon	0.069478	0.254316	0.083711	0.276957	-0.01423	0.0127	0.0073
Dumfar	0.010843	0.103586	0.012034	0.109037	-0.00119	0.5971	0.5805
Dummin	0.006426	0.079919	0.007838	0.088188	-0.00141	0.4368	0.3969
Dumene	0.002811	0.052957	0.002456	0.049502	0.000355	0.7305	0.7454
Dumbus	0.150602	0.357733	0.133915	0.340565	0.016688	0.0184	0.0240
Dumpub	0.018474	0.134684	0.021473	0.144956	-0.003	0.3159	0.2850
Dumtrahot	0.35261	0.477879	0.34544	0.475518	0.00717	0.4669	0.4689
Dumtra	0.051004	0.22005	0.058043	0.233827	-0.00704	0.1448	0.1243

Note: See the note to table 12.2.a.

Italy: 2000 sample

The E5 firms and the non-E5 firms are not as similar in the Italian case as they were in the French case (table 12.2.c). In fact, using the 2000 sample, out of the 22 core, proxy and control variables, 4 core variables, 2 – 3 proxies (depending on the test) and 4 – 6 control variables (depending on the test) are found to have significantly different means.

The E5 firms are younger, bigger, have a smaller earnings ratio and a smaller solvency ratio compared to the non-E5 firms. Both the solvency ratio and the earnings ratio were hypothesized to affect the firms in financial distress negatively, meaning that the larger these ratios are, the less likely a firm is to enter financial distress. As the E5 firms have a solvency ratio and an earnings ratio that is lower than the ratios of the non-E5 firms, on the one hand, one could fear that firms in financial distress are overrepresented in the Italian group of E5 firms. On the other hand, it is comforting that only 2 – 3 (depending on the test) of the 6 proxies have significantly different means. These variables are the number of subsidiaries, the number of shareholders (only Satterthwaite test) and the bvd_indep_b

indicator. No hypothesis is set up on the effect of these variables, and so it is not known in what direction it affects firms in financial distress, and, accordingly, no statements are made on the potential bias of the group of the E5 firms. Furthermore, on the comforting side, E5 firms are younger and bigger. Both age and size were hypothesized to affect the firms in financial distress negatively.

The sector affiliation dummies that have significantly different means between the E5 and the non-E5 firms are: dumman, dumorg (only Satterthwaite test), dumfar (only Satterthwaite test), dumene, dumpub and dumtra (only t-test).

The variables, which cannot be rejected to have equal means, are the loans to total assets ratio, legal form, bvd_indep_a, bvd_indep_c, IT dummy, dumcon, dummin, dumbus and dumtrahot.

Table 12.2.c: Italy: Descriptive Statistics and tests: 2000 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard deviation	Mean	T-test	Satterthwaite
Age	19.53208	14.11504	20.59208	13.47275	-1.06	0.0059	0.0085
Size	8.754034	0.780465	8.687271	0.749118	0.066763	0.0018	0.0027
Earnings ratio	0.074727	0.124645	0.101222	0.095617	-0.0265	<0.0001	<0.0001
Solvency ratio	0.184355	0.218105	0.222743	0.189291	-0.03839	<0.0001	<0.0001
Loans to assets ratio	0.129203	0.159666	0.128073	0.156191	0.00113	0.8000	0.8040
Legal form	0.653365	0.476085	0.643825	0.478875	0.00954	0.4850	0.4827
Shareholders	0.529734	1.245939	0.607024	1.641441	-0.07729	0.0961	0.0322
Subsidiaries	0.374804	0.982011	0.490435	1.275643	-0.11563	0.0014	<0.0001
Bvd_indep_a	0.015649	0.124164	0.015831	0.124822	-0.00018	0.9594	0.9592
Bvd_indep_b	0.028951	0.167736	0.041782	0.200093	-0.01283	0.0238	0.0079
Bvd_indep_c	0.15493	0.361979	0.135751	0.342529	0.019179	0.0503	0.0630
It dummy	0.031299	0.174192	0.022905	0.149604	0.008394	0.0509	0.0899
Dumman	0.507825	0.500134	0.536373	0.498683	-0.02855	0.0449	0.0456
Dumorg	0.007825	0.088145	0.012849	0.112623	-0.00502	0.1152	0.0487
Dumcon	0.08529	0.279421	0.072174	0.25878	0.013115	0.0767	0.0993
Dumfar	0.001565	0.039544	0.004695	0.068362	-0.00313	0.1040	0.0076
Dummin	0.00626	0.078902	0.00717	0.084372	-0.00091	0.7047	0.6869
Dumene	0.010172	0.100382	0.003077	0.055389	0.007095	<0.0001	0.0122
Dumbus	0.047731	0.21328	0.049332	0.216564	-0.0016	0.7954	0.7926
Dumpub	0.00313	0.05588	0.0125	0.111102	-0.00937	0.0027	<0.0001
Dumtrahot	0.277778	0.448079	0.257543	0.437288	0.020235	0.1052	0.1134
Dumtra	0.050861	0.219799	0.039656	0.195153	0.011205	0.0453	0.0730

Note: See the note to table 12.2.a.

Table 12.2.d: Italy: Descriptive Statistics and tests: 2001 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard deviation	Mean	T-test	Satterth-waite test
Age	19.952077	12.811314	20.945696	13.467745	-0.993619	<0.0001	<0.0001
Size	8.646506	0.955063	8.625972	0.809803	0.020534	0.1411	0.1962
Earnings ratio	0.078043	0.138421	0.100377	0.110331	-0.022334	<0.0001	<0.0001
Solvency ratio	0.214996	0.221759	0.227800	0.208256	-0.012804	0.0003	0.0006
Loans to assets ratio	0.126792	0.159462	0.128733	0.161791	-0.001941	0.4767	0.4717
Legal form	0.670405	0.470126	0.672037	0.469478	-0.001631	0.8370	0.8371
Shareholders	0.550599	1.602260	0.578574	1.566138	-0.027975	0.2913	0.3000
Subsidiaries	0.454754	1.332403	0.442725	1.174124	0.012029	0.5500	0.5884
Bvd_indep_a	0.015294	0.122737	0.014843	0.120926	0.000451	0.8253	0.8273
Bvd_indep_b	0.033903	0.181002	0.040758	0.197732	-0.006855	0.0383	0.0264
Bvd_indep_c	0.133826	0.340509	0.130329	0.336670	0.003498	0.5389	0.5425
It dummy	0.029569	0.169417	0.022566	0.148518	0.007003	0.0060	0.0132
Dumman	0.512873	0.499898	0.523607	0.499450	-0.010734	0.2031	0.2035
Dumorg	0.009432	0.096669	0.013516	0.115470	-0.004084	0.0333	0.0144
Dumcon	0.076982	0.266597	0.074547	0.262662	0.002435	0.5835	0.5880
Dumfar	0.006373	0.079584	0.004586	0.067563	0.001787	0.1247	0.1771
Dummin	0.006118	0.077986	0.007361	0.085482	-0.001243	0.3847	0.3502
Dumene	0.005353	0.072978	0.002926	0.054017	0.002427	0.0107	0.0436
Dumbus	0.058884	0.235437	0.050442	0.218858	0.008442	0.0235	0.0325
Dumpub	0.010706	0.102928	0.012339	0.110395	-0.001633	0.3777	0.3513
Dumtrahot	0.264848	0.441309	0.267837	0.442839	-0.002989	0.6892	0.6884
Dumtra	0.039256	0.194227	0.039430	0.194620	-0.000175	0.9576	0.9575

Note: See the note to table 12.2.a.

Italy: 2001 sample

Using the 2001 sample, out of the 22 core, proxy and control variables, 3 core variables, only 1 proxy variable and 4 sector affiliation dummies are rejected have equal means (table 12.2.d). The core variables that are rejected to have equal means are age, the earnings ratio and the solvency ratio. As was the case for the Italian 2000 sample the E5 firms are younger, they have smaller earnings ratios and smaller solvency ratios. Both the solvency ratio and the earnings ratio were hypothesized to affect the firms in financial distress negatively. As the E5 firms have lower ratios than non-E5 firms, one could fear that firms in financial distress may be over-represented in the Italian group of E5 firms. Age was hypothesized to affect the firms in financial distress negatively.

The proxy variable, which is rejected to have equal means, is bvd_indep_b. That this variable is significantly different between E5 and non-E5 firms is not thought to affect the sample in the way, such that firms in financial distress are more likely to be present among the E5 firms (as this variable was hypothesized not to have an effect in the final estimations).

The sector affiliation dummies that are rejected to have equal means are: IT dummy, dumorg, dumene and dumbus.

The variables, which cannot be rejected to have equal means, are the loans to assets ratio, size, legal form, the number of shareholders, the number of subsidiaries, bvd_indep_a, bvd_indep_b, dumman, dumcon, dumfar, dummin, dumpub, dumtrahot and dumtra.

Italy: Conclusion

The main worry concerning the Italian data is that the solvency ratio and the earnings ratio is significantly different (and smaller for the E5 firms) in both the 2000 and the 2001 sample. This could indicate a bias in the data. Firms in financial distress may be more likely to be among the E5 firms. On the other hand, on the positive side, only 2 – 3 (depending on the test) of the 6 proxies have significantly different means in the 2000 sample and 1 of the 6 proxies have significantly different means in the 2001 sample (bvd_indep_b, which is significantly different for E5 and non-E5 firms, was hypothesized not to affect firms in financial distress). Furthermore, on the comforting side, age and size, which were hypothesized to affect the firms in financial distress negatively, affect the firms in opposite directions (2000 sample). In the 2001 sample E5 firms are younger than Non-E5 firms.

The overall conclusion in the Italian case is not as clear as the conclusion in the French case. There seem to be some systematic differences in the Italian dataset. These differences may cause concern about what can be inferred with confidence from the dataset.

Spain: 2000 sample

As for the Italian case, the Spanish E5 firms and the Spanish non-E5 firms are not as similar as they were in the French case (table 13.1.e). Out of the 22 core, proxy and control variables, 9 – 10 variables are rejected to have equal means (depending on the test).

All the 5 core variables are rejected to have equal means. The E5 firms are younger, smaller, have a lower earnings ratio and solvency ratio and their loans to total assets ratio is smaller compared to the non-E5 firms. That the earnings and the solvency ratio are smaller for E5 firms compared to non-E5 firms indicates, that there may be a bias. From looking at these two variables alone, firms in financial distress seem to be more likely to be among the E5 firms. Furthermore, age and size were assumed to affect the firms in financial distress negatively. The opposite conclusion is obtained when the loans to total assets ratio is at focus. This ratio is hypothesized to have a positive sign in the estimations, which would indicate, that firms with a high loans to total assets ratio are more likely to enter financial distress. The E5 firms have a lower ratio compared to the non-E5 firms, and so, based on the loans to total assets variable, weak firms do not seem to be over-represented in the E5 group.

Only 1 – 2 (depending on the test) proxies are rejected to have equal means. These proxies are the number of subsidiaries and the bvd_indep_a (Satterthwaite test only). No hypotheses are set up on either variable.

The following sector affiliation dummies are rejected to have equal means: dumorg, dumbus and dumpub.

The variables, which cannot be rejected to have equal means, are: legal form, number of shareholders, bvd_indep_a (only t-test), bvd_indep_b, bvd_indep_c, IT dummy, dumman, dumcon, dumfar, dummin, dumene, dumtrahot and dumtra.

Table 12.2.e: Spain: Descriptive Statistics and t-tests: 2000 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard Deviation	Mean	T-test	Satterthwaite
Age	14.44123	11.75834	16.76956	11.90877	-2.32832	<0.0001	<0.0001
Size	8.443674	0.685648	8.550479	0.733734	-0.10681	0.0010	0.0005
Earnings ratio	0.08631	0.117609	0.107376	0.113418	-0.02107	<0.0001	<0.0001
Solvency ratio	0.29513	0.28334	0.342124	0.237304	-0.04699	<0.0001	0.0002
Loans to assets ratio	0.051159	0.127056	0.067702	0.124598	-0.01654	0.0028	0.0035
Legal form	0.362235	0.48111	0.337811	0.472974	0.024424	0.2451	0.2533
Shareholders	2.104046	3.158118	2.230489	3.070334	-0.12644	0.3541	0.3674
Subsidiaries	0.554913	1.362901	0.952873	2.427232	-0.39796	0.0002	<0.0001
Bvd_indep_a	0.036609	0.187981	0.055642	0.229233	-0.01903	0.0606	0.0238
Bvd_indep_b	0.123314	0.329115	0.110145	0.313077	0.013169	0.3442	0.3675
Bvd_indep_c	0.375723	0.484776	0.396275	0.489134	-0.02055	0.3440	0.3403
It dummy	0.026975	0.162167	0.016529	0.127499	0.010446	0.0670	0.1457
Dumman	0.339114	0.473865	0.368682	0.482459	-0.02957	0.1674	0.1608
Dumorg	0.050096	0.218354	0.028322	0.165894	0.021775	0.0034	0.0244
Dumcon	0.092486	0.28999	0.107595	0.309875	-0.01511	0.2716	0.2420
Dumfar	0.019268	0.137597	0.020854	0.142899	-0.00159	0.8025	0.7954
Dummin	0.007707	0.087536	0.012749	0.112193	-0.00504	0.3094	0.1985
Dumene	0.00578	0.075882	0.005601	0.074629	0.00018	0.9568	0.9575
Dumbus	0.117534	0.322366	0.088471	0.283985	0.029063	0.0216	0.0423
Dumpub	0.026975	0.162167	0.012704	0.111995	0.014271	0.0046	0.0467
Dumtrahot	0.26975	0.444258	0.292095	0.454736	-0.02235	0.2683	0.2581
Dumtra	0.071291	0.257558	0.062927	0.242837	0.008364	0.4387	0.4644

Note: See the note to table 12.2.a.

Spain: 2001 sample

The Spanish 2001 sample shows a somewhat similar picture (table 13.1.f). Out of the 22 core, proxy and control variables, 8 – 9 variables are rejected to have equal means (depending on the test).

4 out of the 5 core variables are rejected to have equal means. These are age, the earnings ratio, the solvency ratio and the loans to total assets ratio. The E5 firms are younger, have a lower earnings and solvency ratio and a lower loans to total assets ratio compared with the non-E5 firms. That E5 firms have smaller earnings ratios and smaller solvency ratios indicates that there may be a bias towards financially distressed firms being over-represented in the E5 sample. Furthermore, age was

hypothesized to affect firms in financial distress negatively. On the other hand, the loans to total assets ratio, which is hypothesized to affect the Spanish firms in financial distress with a positive sign, is smaller for E5 firms than for non-E5 firms. This would lead to the opposite conclusion, namely that firms in financial distress are not likely to be over-represented among the E5 firms.

Only 1 out of the 6 proxies is significantly different between the E5 and the non-E5 firms. This is the variable number of subsidiaries. The variable was not hypothesized to have a specific effect on the firms in financial distress. It was left to the estimations to show the direction of the effect, if there is one.

3 – 4 sectors have significantly different means (depending on the test): dumman, dumpub (only Satterthwaite test), dumtrahot, dumtra.

The variables, which cannot be rejected to have equal means, are: legal form, shareholders, bvd_indep_a, bvd_indep_b, bvd_indep_c, IT dummy, dumorg, dumcon, dumfar, dummin, dumene, dumbus and dumpub (only Satterthwaite test).

Table 12.2.f: Spain: Descriptive Statistics and t-tests: 2001 sample

	Attritioners		Non-Attritioners		Difference	Tests: $\Pr > \|t\|$	
	Mean	Standard deviation	Mean	Standard deviation	Mean	T-test	Satterthwaite test
Age	15.99909	12.01406	16.72228	11.6478	-0.723192	0.0053	0.0066
Size	8.513385	0.731434	8.522244	0.75797	-0.008859	0.5975	0.5864
Earnings ratio	0.084082	0.123399	0.102988	0.124393	-0.018906	<0.0001	<0.0001
Solvency ratio	0.31409	0.274709	0.344495	0.249143	-0.030405	<0.0001	<0.0001
Loans to assets ratio	0.058141	0.12142	0.064638	0.122947	-0.006497	0.0172	0.0161
Legal form	0.376134	0.484524	0.377789	0.484844	-0.001655	0.8778	0.8777
Shareholders	2.132033	3.099542	2.150812	2.949067	-0.018779	0.7751	0.7841
Subsidiaries	0.746824	1.813943	0.871734	2.316939	-0.124910	0.0136	0.0025
Bvd_indep_a	0.059437	0.236495	0.054126	0.22627	0.005312	0.2920	0.3101
Bvd_indep_b	0.098004	0.297387	0.106655	0.30868	-0.008651	0.2054	0.1915
Bvd_indep_c	0.389292	0.4877	0.386278	0.486905	0.003014	0.7803	0.7806
It dummy	0.019056	0.136754	0.0169	0.128898	0.002157	0.4532	0.4755
Dumman	0.298094	0.457525	0.355048	0.478537	-0.056954	<0.0001	<0.0001
Dumorg	0.03176	0.175401	0.028309	0.165857	0.003452	0.3507	0.3734
Dumcon	0.118875	0.323715	0.116156	0.320418	0.002719	0.7025	0.7049
Dumfar	0.02677	0.161446	0.021456	0.1449	0.005314	0.1017	0.1352
Dummin	0.015426	0.12327	0.011876	0.108332	0.003550	0.1444	0.1905
Dumene	0.007713	0.087506	0.005452	0.073634	0.002262	0.1733	0.2388
Dumbus	0.084846	0.278715	0.096336	0.295057	-0.011490	0.0781	0.0646
Dumpub	0.018603	0.135147	0.013356	0.114797	0.005246	0.0425	0.0771
Dumtrahot	0.322595	0.467575	0.290838	0.454158	0.031758	0.0017	0.0022
Dumtra	0.075318	0.263963	0.061174	0.239653	0.014144	0.0084	0.0151

Note: See the note to table 12.2.a.

Spain: Conclusion

In both the 2000 and the 2001 sample the earnings ratio and the solvency ratio is lower for E5 firms compared to non-E5 firms. Furthermore the E5 firms are younger (2000 and 2001 sample) and smaller (2000 sample) than the non-E5 firms. This could point towards a bias in the data, namely that firms in financial distress are likely to be over-represented among the E5 firms. On the other hand, the loans to total assets ratio, which is rejected to have equal means in both samples, is in both samples lower for the E5 firms compared to the non-E5 firms. Looking at this variable alone, one would conclude that firms in financial distress are not over-represented in the E5 group. Further comforting evidence is obtained when the proxies are compared. In the 2000 sample 1 – 2 proxies (depending on the test) are rejected to have equal means and in the 2001 sample 1 proxy is rejected to have equal means.

The evidence in the Spanish case is mixed. Because of the smaller earnings and solvency ratio in both samples and because of the effects of age and size, it seems as if there is some systematic variation in the Spanish data, and that firms in financial distress may be over-represented in the E5 group. Nonetheless, the loans to total assets ratio points in the other direction, indicating that firms in financial distress are not over-represented in the group of E5 firms. Furthermore only few proxies are found to be significantly different in the samples. If there are systematic differences in the data, it may cause concern about what can be inferred from data.

12.3 The Probability of Attrition

To investigate the firms further, attrition probits, which estimates the probability of attrition, are presented. Probit models are estimated for each country using the 2000 and the 2001 sample, respectively, c.f. tables 12.3.a and 12.3.b. The dependent variable is whether there was attrition or not, and so the sign of the coefficients to the explanatory variables show in what direction the variables affect the likelihood of attrition. χ^2 tests for the significance of the overall relations are presented at the bottom of each table. The null hypothesis is that all coefficients are zero. The alternative hypothesis is that at least one of the coefficients is not 0. The χ^2 tests show, that at least one of the coefficients is not 0 in all countries, both samples.

France: 2000 sample

When estimating the French attrition probit using the 2000 sample, 3 core variables are significant. These are age, size and the solvency ratio. Age has a negative sign, size a positive sign, and the solvency ratio a positive sign. That the solvency ratio has a positive sign indicates that the firms that have a high solvency ratio are more likely to be among the attritioners, i.e. the same result as in the comparison of means (2000 sample). The interpretation of this is that the E5 firms are stronger than the non-E5 firms. Young firms and large firms are more likely to attrite. The comparison of the means

showed that age is rejected to have equal means, but that size could not be rejected to have equal means (Satterthwaite test).

Two proxies are significant in the estimation of the attrition probit: Legal form and number of shareholders. Legal form affects the likelihood of exiting as an E5 firm in a positive way. The comparison of means showed that private limited liability companies are more likely to be among the E5 firms compared to the non-E5 firms, i.e. same conclusion as in the attrition probit. The number of shareholders affects the likelihood of exiting as an E5 firm in a negative way, despite the fact, that the variable could not be rejected to have equal means for E5 and non-E5 firms.

The significant sector affiliation dummies are dumcon and dumpub. In comparison, the sector affiliation dummies that are rejected to have equal means are: dumman, dumcon, dummin (only Satterhwaite) and dumpub.

France: 2001 sample

2 core variables are significant. These are age and the earnings ratio. Both variables have a negative sign. This means that the older the firms are the less likely they are to attrite, and that the higher the earnings ratio is, the less likely they are to attrite. In the comparison of means the same effects were found. The earnings ratio was found to be lower for the E5 firms than for the non-E5 firms, indicating that the E5 firms are weaker than the non-E5 firms, and age was found to be lower for the E5 firms than for the non-E5 firms, indicating that the E5 firms are younger than the non-E5 firms.

1 proxy is significant in the attrition probit. The significant proxy is the variable legal form. It is significant and has a positive sign in the attrition probit, indicating that private limited liability compares are more likely to attrite. The comparison of means showed a similar result. Based on the comparison of means, the conclusion was that private limited liability companies seem to be over-represented among the E5 firms.

There is only 1 significant sector affiliation dummy, namely dumcon. In comparison, the sector affiliation dummies that are rejected to have equal means are dumcon and dumbus.

France: Conclusion

The comparison of the attrition probits with the results obtained in the sections on comparisons of means shows that the results are very alike. The overall conclusion previously obtained, namely that the variables, which are central to the study do not differ in a systematically way between E5 firms and non-E5 firms in the French case, seems to hold.

Table 12.3.a: Attrition probit: 2000 sample

	Spain	Italy	France
Age	-0.00423*	-0.00142	-0.00361*
Size	-0.0342	0.0892*	0.0548*
Earnings ratio	-0.4458*	-0.9130*	-0.1101
Solvency ratio	-0.2208*	-0.2638*	0.2185*
Loans to assets ratio	-0.3525*	-0.2129*	0.1849
Legal form	-0.0331	0.0357	0.1685*
Shareholders	0.000420	-0.0307*	-0.0407*
Subsidiaries	-0.0432*	-0.0498*	-0.00383
Bvd_indep_a	-0.1128	0.1107	0.1284
Bvd_indep_b	0.0607	-0.0455	0.0514
Bvd_indep_c	-0.0334	0.0963*	0.0615
IT dummy	0.0765	0.1954*	0.0562
Dumorg	0.2643*	-0.00142	-0.0877
Dumcon	-0.0764	-0.1921	-0.2173*
Dumfar	-0.0133	0.0494	-0.1252
Dummin	-0.1250	-0.4558	-0.3962
Dumene	0.0724	-0.00468	-0.3808
Dumbus	0.1059	0.6022*	-0.0278
Dumpub	0.3439*	-0.0553	-0.4565*
Dumtrahot	-0.0362	-0.5514*	-0.0696
Dumtra	0.0527	0.0286	-0.0487
Constant	-1.4586*	-2.3604*	-2.3353*
Test of overall significance of model $\Pr > \chi^2$	<0.0001	<0.0001	<0.0001

Note: For more details on the variables the reader is referred to section 3.2. The following abbreviations are used: Dumfar = Farming, forestry and fishing, Dummin = Mining, Dumman= Manufacturing, Dumener = Energy, Dumcon = Construction, Dumtraho = Trade and hotel, Dumtra = Transport, Dumbus = Business service, Dumpub = Public service activities, Dumorg = Organisations. Subsidiaries: This variable measures the number of subsidiaries that a company has registered. Legal form: This dummy is equal to 1, if it is a private limited liability company, and equal to 0, if it is a public limited liability company. Shareholders: This variable measures the number of recorded shareholders. Independence indicators: Bvd_ind_a, bvd_ind_b and bvd_ind_c.

Table 12.3.b: Attrition probit: 2001 sample

	Spain	Italy	France
Age	-0.00149	-0.00299*	-0.00271*
Size	0.0265	0.0265*	-0.00893
Earnings ratio	-0.4326*	-0.8554*	-0.2015*
Solvency ratio	-0.1566*	0.0568	0.0582
Loans to assets ratio	-0.2891*	-0.1308*	0.1184
Legal form	-0.0317	0.00142	0.1692*
Shareholders	0.000766	-0.00478	-0.0160
Subsidiaries	-0.0163	0.00240	0.0115
Bvd_indep_a	0.0674	0.0187	0.1171
Bvd_indep_b	-0.0307	-0.0826	0.0790
Bvd_indep_c	-0.00336	0.000265	0.0406
IT dummy	0.0321	0.0944	-0.0142
Dumorg	0.1250	-0.1757	0.0177
Dumcon	0.0729	-0.00163	-0.1036*
Dumfar	0.1863*	0.1434	-0.0764
Dummin	0.2463*	-0.0626	-0.0945
Dumene	0.2502	0.3389*	0.0966
Dumbus	-0.0149	0.0420	0.0257
Dumpub	0.2351*	-0.0247	-0.0934
Dumtrahot	0.1233*	-0.0166	-0.0103
Dumtra	0.1810*	-0.0034	-0.0822
Constant	-1.5517*	-1.3301*	-1.4194*
Test of overall significance of model $\Pr > \chi^2$	<0.0001	<0.0001	<0.0001

Note: See the note to table 12.3.a.

Italy: 2000 sample

In the Italian case 4 core variables are significant. The solvency ratio, the earnings ratio and the loans to assets ratio is significant and have a negative sign. The size variable is significant and has a positive sign. The comparison of means showed E5 firms could not be rejected to be larger, have smaller earnings ratios and smaller solvency ratios, and so for these three variables the comparison of means and the attrition probit give consistent results. The loans to total assets ratio could not be rejected to have equal means, and so the attrition probit and the univariate comparisons give conflicting results. Based on the effects of the earnings ratio and the solvency, one could argue that firms in financial

distress may be more likely to be among the E5 firms, as these variables were hypothesized to affect firms in financial distress negatively. With size the conclusion is the opposite, as size was hypothesized to affect firms in financial distress negatively.

3 proxies are significant in the estimations: the number of shareholders (-), the number of subsidiaries (-) and the bvd_indep_c (+). The means that were rejected to be equal were the means of the variables: number of shareholders, number of subsidiaries and bvd_indep_b.

The significant sector affiliation dummies are IT dummy, dumbus and dumtrahot. In comparison, the sector affiliation dummies that have significantly different means (E5 and the non-E5 firms) are: dumman, dumorg (only Satterthwaite test), dumfar (only Satterthwaite test), dumene, dumpub and dumtra (only t-test).

Italy: 2001 sample

As for the 2000 sample, 4 core variables are significant in the attrition probit using the 2001 sample. The significant variables in the 2001 sample (which differ from the 2000 sample) are age, size, the earnings ratio and the loans to total assets ratio. Age affects the likelihood of exiting as an E5 firm negatively and size affects the likelihood of exiting as an E5 firm positively. The coefficients to the earnings ratio and the loans to total assets ratio are negative, indicating that a higher earnings ratio and a higher loans to total assets ratio, the less likely it is that the firm is an E5 firm. A high earnings ratio is hypothesized to affect firms in financial distress with a negative sign. That firms with high earnings ratio are more likely to attrite is therefore an indication of them being financially distressed. The loans to total assets ratio is hypothesized to affect firms in financial distress positively. It is noteworthy that the solvency ratio (which, measured on the mean, is smaller for E5 firms compared to non-E5 firms) is insignificant in the attrition probit. The solvency ratio is expected to have a negative sign in the estimations of the credit-scoring model.

The comparison of means showed that age, the earnings ratio and the solvency ratio are rejected to have equal means (2001 sample), and so only age and the earnings ratio are consistent between the comparisons of means and the attrition probit. The solvency ratio is insignificant in the attrition probit, but the mean of the E5 and non-E5 firms is rejected to be equal. Size and the loans to total assets ratio is significant in the attrition probit, but the means cannot be rejected to be equal.

None of the proxy variables are significant in the attrition probit. The only proxy, which is rejected to have equal means, is bvd_indep_b.

Only one sector affiliation dummy is significant: dumene. In comparison, the sector affiliation dummies that are rejected to have equal means are: IT dummy, dumorg, dumene and dumbus.

Italy: Conclusion

The results from the attrition probits show that the effects differ from the 2000 and the 2001 sample, in fact only three core variables have the same effect in both samples, namely size (+), the earnings ratio (-) and the loans to total assets ratio (-). Most importantly the solvency ratio, which is hypothesized to have a positive coefficient in the credit-scoring model, is significant and has a negative sign using the 2000 sample (attrition probit) and insignificant using the 2001 sample (attrition probit), and the loans to total assets ratio is significant and has a negative sign in the Italian attrition probit in both samples. Age is insignificant using the 2000 sample and significant with a negative sign using the 2001 sample.

A comparison of the results from the comparison of means with the results from the attrition probits shows, that some of the core variables that are significant in the attrition probits have opposite signs from what might be expected from the comparisons of the means, suggesting the opposite relation to attrition if there are multivariate controls. In fact this is the case for the solvency ratio using the 2001 sample (it is insignificant in the attrition probit, but it is rejected to have equal means, with E5 firms having lower ratios than non-E5 firms), age using the 2000 sample (it is insignificant in the attrition probit, but it is rejected to have equal means, with E5 firms being younger than non-E5 firms), and the loans to total assets ratios in the 2000 and 2001 sample (it is significant in the attrition probit with a negative sign, but it cannot be rejected to have equal means).

The results from the attrition probits show that the effects of the proxy variables also are not the same in the two years. Using the 2000 sample, 3 proxies are significant (the number of shareholders (-), the number of subsidiaries (-) and the bvd_indep_c (+)). None of the proxies are significant using the 2001 sample. And so there seems to be no consistent tendency over time. Furthermore, there is some discrepancy between the results in the attrition probit and the comparisons of means. The means that were rejected to be equal using the 2000 sample were the means of the variables: number of shareholders, number of subsidiaries and bvd_indep_b. The means that were rejected to be equal using the 2001 sample was the mean of the variable bvd_indep_b.

All in all, firstly, the results of the attrition probits are not consistent over time, and, secondly, the results show that the characteristics, which predict attrition with multivariate controls, and the directions of those effects inferred simply by examining the significance of means in univariate comparisons between the subsamples, lead to opposing results in quite a number of cases. While the comparisons of means suggested that worse-off firms may be more likely to be among the attritioners, the multivariate estimates are less supportive of this conclusion. The key thing to note is that the solvency ratio is not significant in the 2001 sample in the attrition probit.

Spain: 2000 sample

The estimation of the attrition probit shows that age, the earnings ratio, the solvency ratio and the loans to assets ratio are significant predictors of attrition. The earnings ratio and the solvency ratio were

assumed to affect the likelihood of entering financial distress negatively, meaning that the higher these ratios are the less likely the firms are to enter financial distress. The attrition probit shows that these ratios affect attrition the same way as they affect financial distress. The tests of equal means showed as well, that the earnings ratio and the solvency ratio are rejected to have equal means and that the mean of the ratios are smaller for E5 firms than for non-E5 firms. The loans to assets ratio is hypothesized to be significant and have a positive sign in the credit-scoring model. Here the ratio has a negative sign indicating that the higher the ratio is, the less likely the firm is to be an attritioner. The comparison of means suggests the same relation between E5 firms and non-E5 firms. The loans to assets ratio is rejected to have equal means, and E5 firms have a lower loans to assets ratio compared to non-E5 firms. Age is significant with a negative sign in the estimation of the attrition probit. The comparison of means suggests the same relation between the E5 firms and the non-E5 firms, i.e. attrition seems to be more outspoken for younger firms. The only core variable, for which conflicting results are obtained, is the size variable. Size is insignificant in the attrition probit, but the comparisons of means showed that they are rejected to have equal means, and that attritioners were smaller than non-attritioners.

Only 1 proxy is significant in the attrition probit, namely the number of subsidiaries. The E5 firms and the non-E5 firms are rejected to have equal means for this variable. Both results point in the direction of E5 firms having a smaller number of subsidiaries than non-E5 firms. The bvd_indep_a is rejected to have equal means, but is not significant in the attrition probit.

Only two sector affiliation dummies are significant in the attrition probit: Dumorg and Dumpub. In comparison, the following sector affiliation dummies are rejected to have equal means: dumorg, dumbus and dumpub.

Spain: 2001 sample

The attrition probit shows that 3 core variables are significant and all have negative coefficients. These are the earnings ratio, the solvency ratio and the loans to assets ratio. The same variables were significant and had a negative coefficient also in the attrition probit using the 2000 sample. As for the comparison of means in the 2000 sample, the means of these variables are rejected to be equal using the 2001 sample (and all ratios are smaller for the E5 firms compared to the non-E5 firms).

In the 2001 attrition probit age and size are not significant (age was significant in the 2000 attrition probit). This is in contrast to the comparisons of means, which showed that the mean of the age of the companies is rejected to be equal in the 2001 sample (age is smaller for E5 firms). The means of the variable size cannot be rejected to be equal.

None of the proxies are significant in the 2001 attrition probit, whereas the mean of 1 proxy is rejected to be equal for the E5 and non-E5 firms (number of subsidiaries).

5 sector affiliation dummies are significant in the attrition probit: dumfar, dummin, dumpub, dumtrahot and dumtra. In comparison, 3 – 4 sector affiliation dummies have different means (depending on the test): dumman, dumpub (only Satterthwaite test), dumtrahot, dumtra.

Spain: Conclusion

The estimations of the attrition probits show that 3 core variables are significant (with negative coefficients) using both the 2000 and the 2001 sample. These are the earnings ratio, the solvency ratio and the loans to assets ratio. For both samples, the comparison of means of the E5 firms with the non-E5 firms shows that the means of these variables are rejected to be equal using both samples, and that the ratios are smaller for the E5 firms compared to the non-E5 firms. The results on the earnings ratio and the solvency ratio, which are assumed to affect firms in financial distress negatively, are worrying, as they indicate that firms that are worse-off are more likely to be among the attritioners. On the other hand the comparison of the means and the results from the attrition probits based on the loans to assets ratio, which is hypothesized to affect the firms in financial distress positively, is interpreted in the opposite way, i.e. the conclusion from the effects stemming from this variable would be, that firms that are worse-off are less likely to be among the attritioners.

Using the 2000 sample age is significant with a negative sign in the estimation of the attrition probit (the comparison of means suggests the same relation between the E5 firms and the non-E5 firms, i.e. attrition seems to be more outspoken for younger firms), whereas age is insignificant using the 2001 sample (this is in contrast to the comparisons of means, which showed that the mean of the age of the companies is rejected to be equal in the 2001 sample (age is smaller for E5 firms)). Size is insignificant in the attrition probit using the 2000 and the 2001 sample. The comparisons of means showed that size is rejected to have equal means in the 2000 sample (attritioners are smaller than non-attritioners), whereas size could not be rejected to have equal means using the 2001 sample. For the variables age and size there seem to be no clear direction of the results, as some of the significant coefficient estimates is opposite in sign from what might be expected from the comparisons of means or vice versa. This suggests that a different relation to attrition exists if there are multivariate controls included in the estimations.

In the 2000 sample only 1 proxy is significant in the attrition probit, namely the number of subsidiaries (the E5 firms and the non-E5 firms are rejected to have equal means for this variable), whereas in the 2001 sample none of the proxies are significant. In the 2000 sample, a part from the number of subsidiaries, the bvd_indep_a is rejected to have equal means. In the 2001 sample, 1 proxy is rejected to have equal means for the E5 and non-E5 firms (number of subsidiaries).

The overall result is that the sample of Spanish E5 firms may be biased, but not necessarily so. The results go both ways. There are indications that firms that are worse-off are among the attritioners (based on the solvency and the earnings ratio). On the other hand the results on the loans to assets ratio

indicate that the E5 firms are better off. The results coming from the comparisons of the means and the attrition probits on age, size and the proxies are conflicting, indicating that there is no clear pattern in the potential bias of the E5 firms.

12.4 Robustness check

As a robustness check, the credit-scoring models[31] are estimated using 1) a dataset with the E5 firms and with E5 firms as exits, even though we do not know if they are "real" exits, and using 2) a dataset without the E5 firms and with no E5 exit option, and, thereafter, the sign of the coefficients as well as the models predictive ability is compared. Firstly, it is noticed that the variables that are significant when the E5 firms are excluded and when the E5 firms are included as an exit option are the same for the E1 hazards in all countries and with the same sign in all cases, but the legal status variable for Italy. It is no longer significant at the 5 pct. level when the E5 firms are included as an exit option, but it is instead significant at the 9 pct. level. In both specifications the sign is positive. Secondly, a comparison of tables 5.2.a, 5.2.b and 5.2.c with tables 12.4.a, 12.4.b and 12.4.c show that the specifications generate predictions that are very similar.[32]

Table 12.4.a: Competing-risks model with E5 firms: Spain

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)
Event (E1 = financial distress)	Correct call of event: 73 pct. (131 out of 180)	Type 1 error: Missing prediction: 27 pct. (49 out of 180)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)	Type 2 error: Wrong signal: 33 pct. (24,714 out of 79,980)	Correct call of non-event: 67 pct. (53,818 out of 79,980)

[31] See section 4 and 5 for discussions of the methodology and the results, respectively.

[32] As a larger amount of firms are included in the estimations, which include E5 firms, other cut-off levels are used in this case: 1) For France, a cut-off level of 0.01524 (=1,703/111,717) is used. 2) For Spain, a cut-off level of 0.00225 (=180/79,980) is used. 3) For Italy, a cut-off level of 0.00146 (=155/106,011) is used.

Table 12.4.b: Competing-risks model with E5 firms: Italy

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)
Event (E1 = financial distress)	Correct call of event: 76 pct. (118 out of 155)	Type 1 error: Missing prediction: 24 pct. (37 out of 155)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)	Type 2 error: Wrong signal: 27 pct. (28,934 out of 106,011)	Correct call of non-event: 73 pct. (77,077 out of 106,011)

Table 12.4.c: Competing-risks model with E5 firms: France

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)
Event (E1 = financial distress)	Correct call of event: 74 pct. (1,253 out of 1,703)	Type 1 error: Missing prediction: 26 pct. (450 out of 1,703)
Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm, E5 = unknown)	Type 2 error: Wrong signal: 30 pct. (33,299 out of 111,717)	Correct call of non-event: 70 pct. (78,418 out of 111,717)

13. Appendix: Descriptive Statistics

This appendix gives an overview of the data used in the estimations presented in this paper. The definitions of the explanatory variables as well as the abbreviations used are seen in table 3.2 in section 3.2.

Spain

Table 13.1.a: Spain: Number of E1, E2, E3, E4 and active firms

	2000	2001	2002	Total
E1	65	64	51	180
E2	31	41	23	95
E3	272	301	344	917
E4	8	15	27	50
Active	20015	25260	29349	74624
Total	20391	25681	29794	75866

Table 13.1.b: Age and size

		Age			Size		
Exit	Number of firms	average	std. dev.	median	Average	std. dev.	median
E1	180	16.2889	13.8380	13.0000	8.4502	0.7847	8.3473
E2	95	18.3579	14.5748	15.0000	8.7113	0.7689	8.6070
E3	917	16.0862	11.8491	14.0000	8.6708	0.8318	8.5901
E4	50	17.1000	17.4803	12.5000	8.2873	1.0563	8.0609
Active	74624	16.7981	11.6487	15.0000	8.5207	0.7581	8.3584

Table 13.1.c: Earnings ratio, solvency ratio and loans to total assets

		Earnings ratio			Solvency ratio			Loans to total assets		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median	average	std. dev.	median
E1	180	-0.0592	0.3133	0.0330	0.0216	0.4087	0.0947	0.1021	0.1676	0
E2	95	0.0840	0.1604	0.0809	0.2537	0.3390	0.2800	0.0780	0.1294	0.0002
E3	917	0.0643	0.2265	0.0750	0.3108	0.3409	0.3092	0.0577	0.1130	0
E4	50	-0.0153	0.2196	0.0170	0.2536	0.3652	0.2162	0.0548	0.1119	0
Active	74624	0.1027	0.1193	0.0928	0.3464	0.2690	0.3102	0.0641	0.1221	0

Table 13.1.d: Legal form

		Legal form		
Exit	Number of firms	average	std. dev.	median
E1	180	0.4056	0.4924	0.0000
E2	95	0.3368	0.4751	0.0000
E3	917	0.4035	0.4909	0.0000
E4	50	0.4000	0.4949	0.0000
Active	74624	0.3794	0.4852	0.0000

Table 13.1.e: Shareholders and subsidiaries

		Shareholders			Subsidiaries		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	180	2.6278	5.4373	2.0000	0.7611	1.8860	0.0000
E2	95	0.1789	0.5048	0.0000	0.1158	0.8488	0.0000
E3	917	0.5027	1.3953	0.0000	0.1919	1.2033	0.0000
E4	50	2.3200	3.3772	2.0000	0.7800	2.0333	0.0000
Active	74624	2.1593	2.9383	2.0000	0.8708	2.3062	0.0000

Table 13.1.f: Independence indicators

		Bvd_ind_a			Bvd_ind_b			Bvd_ind_c		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median	average	std. dev.	median
E1	180	0.0722	0.2596	0.0000	0.1333	0.3409	0.0000	0.3500	0.4783	0.0000
E2	95	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1158	0.3217	0.0000
E3	917	0.0076	0.0871	0.0000	0.0229	0.1497	0.0000	0.1625	0.3691	0.0000
E4	50	0.1000	0.3030	0.0000	0.1400	0.3505	0.0000	0.4200	0.4986	0.0000
Active	74624	0.0536	0.2252	0.0000	0.1072	0.3094	0.0000	0.3870	0.4871	0.0000

Table 13.1.g: IT and tele dummy

		IT and tele dummy		
Exit	Number of firms	average	std. dev.	Median
E1	180	0.0278	0.1648	0.0000
E2	95	0.0526	0.2245	0.0000
E3	917	0.0676	0.2512	0.0000
E4	50	0.0400	0.1979	0.0000
Active	74624	0.0156	0.1238	0.0000

Table 13.1.h: Sector affiliation

		Mean				
Exit	Number of firms	Dumman	Dumorg	Dumcon	Dumfar	Dummin
E1	180	0.0378	0.0167	0.1667	0.0111	0.0056
E2	95	0.2842	0.0211	0.0105	0.0000	0.0000
E3	917	0.2923	0.0632	0.0513	0.0218	0.0120
E4	50	0.2800	0.0200	0.1200	0.0000	0.0000
Active	74624	0.3559	0.0278	0.1191	0.0214	0.0120

Table 13.1.i: Sector affiliation

		Mean				
Exit	Number of firms	Dumene	Dumbus	Dumpub	Dumtraho	Dumtra
E1	180	0.0000	0.0667	0.0278	0.2389	0.0889
E2	95	0.0105	0.1684	0.0000	0.3895	0.1158
E3	917	0.0120	0.1450	0.0153	0.2868	0.1003
E4	50	0.0000	0.0800	0.0000	0.4000	0.1000
Active	74624	0.0051	0.0945	0.0135	0.2902	0.0606

France

Table 13.2.a: France: Number of E1, E2, E3, E4 and active firms

	2000	2001	2002	Total
E1	167	396	1140	1703
E2	48	686	675	1409
E3	2	22	39	63
E4	152	418	525	1095
Active	28757	34710	40796	104263
Total	29116	36232	43175	108533

Table 13.2.b: Age and size

		Age			Size		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	1703	22.2519	19.8680	16.0000	8.2535	0.6854	8.1236
E2	1409	22.0135	18.0559	17.0000	8.5188	0.7725	8.3705
E3	63	17.3968	14.1928	14.0000	8.0145	0.8903	8.0236
E4	1095	19.3434	17.1048	14.0000	8.3572	0.7533	8.2488
Active	104263	22.4788	17.9819	18.0000	8.4849	0.7303	8.3226

Table 13.2.c. (part 1): Earnings ratio and solvency ratio

		Earnings ratio			Solvency ratio		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	1703	-0.0230	0.2305	0.0292	0.0585	0.3672	0.1121
E2	1409	0.0830	0.1580	0.0793	0.2586	0.3279	0.2584
E3	63	-0.0981	0.4989	-0.0494	-0.1278	0.8616	0.0757
E4	1095	-0.0184	0.2783	0.0333	0.1203	0.4181	0.1424
Active	104263	0.1020	0.1438	0.0962	0.3077	0.2574	0.2937

Table 13.2.c. (part 2): Loans to total assets ratio

		Loans to total assets		
Exit	Number of firms	average	std. dev.	median
E1	1703	0.1117	0.1585	0.0543
E2	1409	0.0743	0.1335	0.0220
E3	63	0.1205	0.2305	0.0031
E4	1095	0.1088	0.1699	0.0389
Active	104263	0.0731	0.1153	0.0307

Table 13.2.d: Legal form

		Legal form		
Exit	Number of firms	average	std. dev.	median
E1	1703	0.1632	0.3697	0.0000
E2	1409	0.1448	0.3520	0.0000
E3	63	0.1587	0.3684	0.0000
E4	1095	0.1489	0.3561	0.0000
Active	104263	0.1372	0.3441	0.0000

Table 13.2.e: Shareholders and subsidiaries

		Shareholders			Subsidiaries		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	1703	1.3030	1.4183	1.0000	0.3958	1.0441	0.0000
E2	1409	0.3385	0.7768	0.0000	0.0298	0.2365	0.0000
E3	63	0.8254	1.4204	0.0000	0.0794	0.3263	0.0000
E4	1095	0.5982	1.0652	0.0000	0.0758	0.4612	0.0000
Active	104263	1.3600	1.2720	1.0000	0.4583	1.2622	0.0000

Table 13.2.f: Independence indicator

		Bvd_ind_a			Bvd_ind_b			Bvd_ind_c		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median	average	std. dev.	median
E1	1703	0.0411	0.1986	0.0000	0.0364	0.1874	0.0000	0.6312	0.4826	1.0000
E2	1409	0.0185	0.1346	0.0000	0.0106	0.1027	0.0000	0.1505	0.3576	0.0000
E3	63	0.0476	0.2147	0.0000	0.0476	0.2147	0.0000	0.2381	0.4293	0.0000
E4	1095	0.0228	0.1494	0.0000	0.0256	0.1579	0.0000	0.2785	0.4485	0.0000
Active	104263	0.0297	0.1697	0.0000	0.0318	0.1754	0.0000	0.6820	0.4657	1.0000

Table 13.2.g: IT and tele dummy

		IT and tele dummy		
Exit	Number of firms	average	std. dev.	median
E1	1703	0.0464	0.2104	0.0000
E2	1409	0.0405	0.1971	0.0000
E3	63	0.0794	0.2725	0.0000
E4	1095	0.0922	0.2895	0.0000
Active	104263	0.0307	0.1726	0.0000

Table 13.2.h: Sector affiliation

		Mean				
Exit	Number of firms	Dumman	Dumorg	Dumcon	Dumfar	Dummin
E1	1703	0.5079	0.0088	0.1133	0.0082	0.0029
E2	1409	0.2576	0.0234	0.0625	0.0099	0.0071
E3	63	0.3333	0.0000	0.0317	0.0159	0.0159
E4	1095	0.3306	0.0192	0.0648	0.0064	0.0055
Active	104263	0.3083	0.0222	0.0843	0.0123	0.0081

Table 13.2.i: Sector affiliation

		Mean				
Exit	Number of firms	Dumene	Dumbus	Dumpub	Dumtraho	Dumtra
E1	1703	0.0000	0.0992	0.0123	0.1908	0.0564
E2	1409	0.0014	0.2229	0.0213	0.3329	0.0610
E3	63	0.0000	0.1905	0.0000	0.3492	0.0635
E4	1095	0.0018	0.2210	0.0110	0.2858	0.0539
Active	104263	0.0026	0.1320	0.0226	0.3494	0.0582

Italy

Table 13.3.a: Italy: Number of E1, E2, E3, E4 and active firms

	2000	2001	2002	Total
E1	73	72	10	155
E2	6	2	16	24
E3	117	74	248	439
E4	9	10	46	65
Active	28083	32989	35977	97049
Total	28288	33147	36297	97732

Table 13.3.b: Age and size

		Age			Size		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	155	15.32258	10.26084	14	8.516675	0.751624	8.5368
E2	24	17.875	11.2687	18	8.835247	1.613305	8.885364
E3	439	20.96128	12.89192	18	8.945507	0.827272	8.930891
E4	65	11.33846	8.522538	9	8.756897	1.321122	8.823501
Active	97049	21.08392	13.72118	19	8.629584	0.812911	8.514991

Table 13.3.c (part 1): Earnings ratio and solvency ratio

		Earnings ratio			Solvency ratio		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median
E1	155	-0.1685	0.3272	-0.0593	-0.1892	0.4706	-0.0127
E2	24	0.0712	0.0771	0.0762	0.2070	0.2359	0.1447
E3	439	0.0605	0.1868	0.0742	0.1817	0.3590	0.1744
E4	65	0.0703	0.0795	0.0671	0.2135	0.2465	0.1304
Active	97049	0.0979	0.1090	0.0882	0.2286	0.2084	0.1848

Table 13.3.c. (part 2): Loans to total assets

		Loans to total assets		
Exit	Number of firms	average	std. dev.	median
E1	155	0.2065	0.2763	0.0967
E2	24	0.1881	0.1950	0.1707
E3	439	0.1581	0.2372	0.0738
E4	65	0.1124	0.1589	0.0022
Active	97049	0.1246	0.1578	0.0476

Table 13.3.d: Legal form

		Legal form		
Exit	Number of firms	Average	std. dev.	median
E1	155	0.7290	0.4459	1.0000
E2	24	0.7083	0.4643	1.0000
E3	439	0.5330	0.4995	1.0000
E4	65	0.7231	0.4510	1.0000
Active	97049	0.6714	0.4697	1.0000

Table 13.3.e: Shareholders and subsidiaries

		Shareholders			Subsidiaries		
Exit	Number of firms	average	std. dev.	median	Average	std. dev.	median
E1	155	0.5677	1.3437	0.0000	0.3290	1.2280	0.0000
E2	24	1.5417	1.8877	2.0000	0.2917	0.6903	0.0000
E3	439	0.7722	3.0689	0.0000	0.7631	1.8563	0.0000
E4	65	0.9385	2.0907	0.0000	0.2154	0.5726	0.0000
Active	97049	0.5776	1.5542	0.0000	0.4443	1.1794	0.0000

Table 13.3.f: Independence indicator

		Bvd_ind_a			Bvd_ind_b			Bvd_ind_c		
Exit	Number of firms	average	std. dev.	median	average	std. dev.	median	average	std. dev.	median
E1	155	0.0129	0.1132	0.0000	0.0194	0.1382	0.0000	0.1613	0.3690	0.0000
E2	24	0.0417	0.2041	0.0000	0.0417	0.2041	0.0000	0.4583	0.5090	0.0000
E3	439	0.0205	0.1419	0.0000	0.0501	0.2184	0.0000	0.1572	0.3644	0.0000
E4	65	0.0154	0.1240	0.0000	0.0615	0.2422	0.0000	0.2000	0.4031	0.0000
Active	97049	0.0148	0.1206	0.0000	0.0407	0.1975	0.0000	0.1299	0.3362	0.0000

Table 13.3.g: IT and tele dummy

		IT and tele dummy		
Exit	Number of firms	Average	std. dev.	Median
E1	155	0.0645	0.2465	0.0000
E2	24	0.0833	0.2823	0.0000
E3	439	0.0364	0.1876	0.0000
E4	65	0.0154	0.1240	0.0000
Active	97049	0.0226	0.1486	0.0000

Table 13.3.h: Sector affiliation

		Mean				
Exit	Number of firms	Dumman	Dumorg	Dumcon	Dumfar	Dummin
E1	155	0.5677	0.0065	0.1355	0.0065	0.0000
E2	24	0.6667	0.0000	0.0000	0.0000	0.0000
E3	439	0.5034	0.0137	0.0569	0.0023	0.0068
E4	65	0.4769	0.0000	0.0769	0.0308	0.0000
Active	97049	0.5236	0.0136	0.0747	0.0046	0.0074

Table 13.3.i: Sector affiliation

		Mean				
Exit	Number of firms	Dumene	Dumbus	Dumpub	Dumtraho	Dumtra
E1	155	0.0000	0.0710	0.0065	0.1677	0.0387
E2	24	0.0000	0.0000	0.0000	0.2917	0.0417
E3	439	0.0046	0.0478	0.0137	0.3007	0.0478
E4	65	0.0154	0.0462	0.0000	0.3385	0.0154
Active	97049	0.0030	0.0511	0.0126	0.2663	0.0396

14. Appendix: The Construction of Sector Affiliation Codes

Table 14.a: Sector affiliation codes

Sector Affiliation	NACE-codes
1. Farming	01
2. Forestry	02
3. Fishing	05
4. Mining	10-14
5. Manufacturing	15-37
6. Energy ("Production of electricity, manufacturing of gas, collection, purification and distribution of water")	40-41
7. Construction ("Construction of buildings and civil engineering works, various contractors and other building completion")	45
8. Trade and hotel ("Wholesale, retail, repair and hotels")	50-52, 55
9. Transport	60-64
10. Business service ("Development and selling of real estate, renting, legal activities, advertising, etc."), (except 74.15: non-financial holding companies)	70-74
11. Public service activities ("General (overall) public service activities, education, hospital activities")	75, 80, 85
12. Organisations, etc. ("Collection and treatment of waste, activities of business and employers organisations, etc., motion picture, video, radio, television, etc., laundering for industrial or commercial clients")	90-93

Table 14.b: The construction of the IT and tele-dummy

NACE codes included in the IT and tele-dummy	Description of sector
3000 – 3100	Manufacture of office machinery and computers
3200	Manufacture of radio, television and communication equipment and apparatus
3220 – 3230	Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy Manufacture of television and radio receivers, sound or video recording or reproducing apparatus and associated goods
6420 – 6421	Telecommunications
7200 – 7300	Computer and related activities Hardware consultancy Software consultancy and supply Data processing

15. Appendix: Results

Tables 15.a, 15.b and 15.c show the parameter estimates for France, Italy and Spain. Tables 15.d, 15.e and 15.f present a detailed breakdown of the discriminatory power in the countries.

The fact that the parameters must be interpreted as contrasts between pairs means, for example, that the odds that an Italian private limited liability company will enter financial distress rather than staying active are about exp(0.4190)=1.52 the odds for Italian public limited liability companies. For the solvency ratio which in the French case has exp(-2.3606)=0.09, each 1-level increase in the variable multiplies the odds of moving into financial distress versus staying active by about 0.09.

The global tests (wald tests, not reported) for the effect of each variable on the outcome variable, controlling for the other variables in the models, shows that in the 1) French case none of the core variables and none of the proxies have no effect on the outcome variable, that in 2) Italian case one of the core variables (loans to total assets ratio) and three of the proxies (number of recorded shareholders, bvd_indep_a and bvd_indep_b) have no effect on the outcome variable, and that in the 3) Spanish case none of the core variables and two of the proxies (legal form and bvd_indep_b) have no effect on the outcome variable.

Table 15.a: Results for France

	E1	E2	E3	E4
Age	0.00302*	0.00212	-0.00745	-0.00308
Size	-0.3265*	0.5722*	-0.4964*	0.1710*
Earnings ratio	-1.6054*	-0.4369*	-1.15674*	-1.8075*
Solvency ratio	-1.5832*	-0.8419*	-1.9564*	-1.1897*
Loans to total assets	0.3169*	-0.5846*	-0.0394	0.4944*
Legal form	0.0624	-0.2267*	-0.3305	-0.2054*
Shareholders	-0.0214	-0.3424*	0.0862	-0.2355*
Subsidiaries	0.0355	-1.8568*	-0.8478*	-0.9966*
Bvd_indep_a	0.1973	-1.1061*	-0.6731	-0.9097*
Bvd_indep_b	-0.0924	-1.3081*	-0.7056	-0.4851*
Bvd_indep_c	-0.3653*	-2.3551*	-2.2554*	-1.6785*

Note: A significance level of 5 pct. is chosen. Year dummies and sector affiliation dummies were included. Because the data was too sparse otherwise in some countries, a grouping of the sector affiliation dummies took place. The sector affiliation dummies that were included in the estimations were: Dumorgpub = dumorg + dumpub, dumcon, dumfarminene = dumfar + dummin + dumene, dumbus, dumtrahot, dumtra. France: In the first estimations age dummies were included. None of these turned out to be significant. In the final estimations no age dummies are included, only the variable age. Italy and Spain: In the first estimations a flexible baseline-hazard function was specified. This led to a quasi-complete separation of data points, meaning that a maximum likelihood estimate may not be possible to obtain, as the data is too sparse. The consequence of this was to use age in the estimations.

Table 15.b: Results for Italy

	E1	E2	E3	E4
Age	-0.0298*	-0.0164	-0.00493	-0.1149*
Size	0.2978*	0.2117	0.3954*	0.3110
Earnings ratio	-2.4314*	-1.2575	-1.3091*	-1.3095
Solvency ratio	-2.3606*	0.6237	-0.6816*	0.7455
Loans to total assets	-0.2008	1.9005	0.1065	-0.4038
Legal form	0.4787*	0.5745	-0.3297*	0.0748
Shareholders	0.0258	0.0437	0.0269	0.0562
Subsidiaries	-0.0954	-0.2823	0.0804*	-0.2721
Bvd_indep_a	-0.0391	1.5015	-0.0503	-0.1619
Bvd_indep_b	-0.2694	0.2656	-0.0201	0.2981
Bvd_indep_c	0.0127	1.6508*	-0.0817	0.2368

Note: See the note to table 15.a.

Table 15.c: Results for Spain

	E1	E2	E3	E4
Age	0.00784	0.0218*	0.00627*	0.00967
Size	0.0294	0.7983*	0.7497*	-0.3548
Earnings ratio	-1.1842*	0.0513	-1.4342*	-2.0180*
Solvency ratio	-0.9511*	-1.5867*	-0.4001*	-0.2487
Loans to total assets	1.4822*	0.6268	0.0570	-0.5464
Legal form	0.1166	-0.3835	-0.0971	0.0524
Shareholders	0.0271	-1.9472*	-0.7760*	-0.0147
Subsidiaries	-0.0516	-0.8504*	-0.5788*	0.00860
Bvd_indep_a	0.1721	-11.7988	-1.5824*	1.1439*
Bvd_indep_b	0.0223	-9.9551	-0.4396	0.6974
Bvd_indep_c	-0.3174	-0.1640	-1.0065*	0.4729

Note: See the note to table 15.a.

Table 15.d: Competing-risks model, detailed breakdown: France

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
E1 (financial distress)	Correct call of event: 75 pct. (1,280 out of 1,703)	Type 1 error: Missing prediction: 25 pct. (423 out of 1,703)
E2 (merger)	Type 2 error: Wrong signal: 39 pct. (550 out of 1,409)	Correct call of non-event: 61 pct. (859 out of 1,409)
E3 (voluntary liquidation)	Type 2 error: Wrong signal: 70 pct. (44 out of 63)	Correct call of non-event: 30 pct. (19 out of 63)
E4 (inactive (no precision))	Type 2 error: Wrong signal: 51 pct. (560 out of 1,095)	Correct call of non-event: 49 pct. (535 out of 1,095)
Active firms	Type 2 error: Wrong signal: 28 pct. (29,347 out of 104,263)	Correct call of non-event: 73 pct. (76,070 out of 104,263)

Table 15.e: Competing-risks model, detailed breakdown: Spain

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
E1 (financial distress)	Correct call of event: 76 pct. (137 out of 180)	Type 1 error: Missing prediction: 24 pct. (43 out of 180)
E2 (merger)	Type 2 error: Wrong signal: 45 pct. (43 out of 95)	Correct call of non-event: 55 pct. (52 out of 95)
E3 (voluntary liquidation)	Type 2 error: Wrong signal: 36 pct. (330 out of 917)	Correct call of non-event: 64 pct. (587 out of 917)
E4 (inactive (no precision))	Type 2 error: Wrong signal: 48 pct. (24 out of 50)	Correct call of non-event: 52 pct. (26 out of 50)
Active firms	Type 2 error: Wrong signal: 32 pct. (23,717 out of 74,624)	Correct call of non-event: 67 pct. (50,907 out of 74,624)

Table 15.f: Competing-risks model, detailed breakdown: Italy

	Model prediction: Event (E1 = financial distress)	Model prediction: Non-event (E2 = merger, E3 = voluntary liquidation, E4 = inactive (no precision) or active firm)
E1 (financial distress)	Correct call of event: 88 pct. (137 out of 155)	Type 1 error: Missing prediction: 12 pct. (18 out of 155)
E2 (merger)	Type 2 error: Wrong signal: 25 pct. (6 out of 24)	Correct call of non-event: 75 pct. (18 out of 24)
E3 (voluntary liquidation)	Type 2 error: Wrong signal: 26 pct. (116 out of 439)	Correct call of non-event: 74 pct. (323 out of 439)
E4 (inactive (no precision))	Type 2 error: Wrong signal: 25 pct. (16 out of 65)	Correct call of non-event: 75 pct. (49 out of 65)
Active firms	Type 2 error: Wrong signal: 29 pct. (28,145 out of 97,049)	Correct call of non-event: 71 pct. (68,904 out of 97,049)