PhD-FDEF-2015-10

The Faculty of Law, Economics and Finance

DISSERTATION

Defense held on 13/07/2015 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN *Sciences Financières*

by

Dimitra MICHALA

Born on 23 October 1984 in Athens, Greece

ESSAYS ON BENCHMARKING CREDIT PERFORMANCE

Dissertation defense committee

Dr Theoharry Grammatikos, dissertation supervisor
Professor, Université du Luxembourg

Dr Jos Van Bommel, Chairman
Professor, Université du Luxembourg

Dr Jean-Daniel Guigou, Vice Chairman
Professor, Université du Luxembourg

Dr Maria-Luisa Ferreira
Economic Advisor, European Investment Bank

Dr Federico Galizia
Economist, Inter-American Development Bank

Abstract

Essays on Benchmarking Credit Performance

By DIMITRA MICHALA

Dissertation Supervisor: Professor Theoharry Grammatikos

The first essay examines idiosyncratic and systematic distress predictors for small and medium sized enterprises (SMEs) using a dataset from eight European countries over the period 2000-2009. In the European Union, small and medium sized enterprises (SMEs) represent 99% of all businesses and contribute to more than half of total value-added. In this essay, we find that SMEs across Europe are vulnerable to common idiosyncratic factors but the relevant systematic factors vary across regions. This indicates the superiority of regional distress models. We also find that systematic factors move average distress rates in the economy and that small SMEs are more vulnerable to these factors compared to large SMEs. By including many very small companies in the sample, this essay offers unique insights into the European small business sector. By exploring distress in a multi-country setting, our models uncover regional vulnerabilities and allow for regional comparisons. Finally, by incorporating systematic dependencies, they capture distress co-movements and clustering.

The second essay provides an explanation of the default anomaly documented in the empirical asset pricing literature. While empirical literature has documented a negative relation between default risk and stock returns, theory suggests that default risk should be positively priced. In this essay, we calculate monthly probabilities of default (PDs) for a large sample of European firms and break them down into systematic and idiosyncratic components. The

approach that we follow does not require data on credit spreads, thus it can also be applied to small firms that do not have such data available. In accordance with theory, we find that the systematic part, measured as the PD sensitivity to aggregate default risk, is positively related to stock returns. We show that stocks with higher PDs underperform because they have, on average, lower exposure to aggregate default risk. Moreover, their idiosyncratic risk is a hedge against downside market conditions. Finally, small and value stocks are quite heterogeneous with respect to such exposure.

The third essay compares private equity-backed IPOs with IPOs of stand-alone companies in a matching framework. The literature suggests that the IPO market may involve higher information asymmetries than acquisitions. Such a setting can influence the behavior of private equity (PE) sponsors as professional insiders. In most cases, I do not find significant differences between these IPOs and comparable IPOs of stand-alone companies. The financial situation of PE-backed companies in the pre-IPO year is similar to that of their peers. PE sponsors do not target their IPOs in "hot" periods any more than would managers of stand-alone companies, nor are they more prone to rush their companies into premature IPOs. They also do not inflate valuations and are not more likely to seek to sell firms with poor prospects ("unload lemons") onto the market. Finally, I find that IPOs that take place in hot periods are significantly more likely to delist due to default, but this result is not any stronger for PE-backed IPOs. This essay provides evidence to contradict media criticism of PE sponsors. This can also have important policy implications regarding the PE regulatory framework related to PE. This work comes as a timely contribution given the increasing importance of PE in the IPO market.

JEL Codes: C13, C41, C53, G11, G12, G15, G23, G24, G32; G33

Keywords: buyout; credit risk; default risk; default anomaly; distress; forecasting; idiosyncratic risk; IPO; logit; Merton model; private equity; SMEs; venture capital

Acknowledgements

The years of my PhD have been the most exciting of my life so far. Many people helped me during this journey on which I gained much valuable experience, met a range of interesting people and traveled widely.

I am particularly grateful to my supervisor, Professor Theoharry Grammatikos for his continuous help and support. He introduced me to the fulfilling world of academic research, and was always available to give me valuable advice.

I am also thankful to my Committee Members, Professor Jos Van Bommel, Professor Jean-Daniel Guigou, Federico Galizia and Maria Luisa Ferreira.

I appreciate enormously the financial support of the European Investment Bank and the help of all EIB staff involved in the process, particularly Maria Luisa Ferreira, Federico Galizia, Edward Olding and Helmut Kraemer-Eis.

My special thanks go to Sara Ferreira Filipe for the great collaboration we shared and for her special help with Chapters 2 and 3.

I would also like to thank Mounir Shal for help with the database used in Chapter 2 and Ali Nasser Eddine for his excellent research assistance in Chapter 3.

Chapter 4 was completed while I was a Visiting Scholar at New York University (Stern School of Business). I am indebted to Professor Christian Wolf and Professor Ingo Walter for making this visit possible, as well as Professor Edward Altman for his valuable feedback while I was at Stern.

I am also grateful to Jay Ritter and Jerry Cao for sharing with me a great deal of the data that I used in Chapter 4.

To my classmates, thank you for your friendship and the moments we shared during these unforgettable years.

Finally, I am eternally thankful to my husband and my parents for their never-ending encouragement and love.

Dedication

To My Husband and My Parents

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Identifying the determinants of corporate credit performance and measuring it correctly is of crucial importance for banks, investors and policy makers. This work focuses on default risk in particular. It explores default risk in private firms (SMEs), listed firms (especially mid and small-caps), and companies that are in the transition phase between private and public ownership (IPOs). The work contributes to the existing literature by (i) uncovering regional corporate default risk vulnerabilities, (ii) explaining the observed anomalous pricing of default risk in the stock market, and (iii) exploring the impact of alternative investment funds such as private equity and venture capital on the default risk of their portfolio companies. In the paragraphs that follow, a summary of the chapters together with the main findings is presented.

Chapter 2 explores the performance of distress prediction hazard models for non-financial SMEs using a dataset from eight European countries over the ten-year period 2000-2009. The panel structure of the dataset allows us to: exploit both the time-series and the cross-sectional dimension; and differentiate between firm-specific, macroeconomic and industry effects.

We find that, in addition to financial indicators (whose importance is noted in past studies) the location and number of shareholders are important determinants of SMEs' distress probabilities. We validate the superiority of models that incorporate macroeconomic dependencies (as suggested by previous research) also in the case of SMEs. However, we do not find strong evidence that industry-effects improve prediction accuracy significantly. We also examine interaction effects between the size of SMEs and systematic variables. We find that as SMEs become larger, they are less vulnerable to macroeconomic changes, contrary to the assumptions inherent in the Basel regulations. Interestingly, when we split our sample into regional sub-samples, we show that SMEs across Europe are sensitive to the same firm-specific factors, but we identify significant regional variations in the selection and importance of macro variables. Specifically, macro variables differ among European regions based on region-specific conditions and characteristics. Since our regional distress models always perform better than a generic model estimated for the regional sub-samples, we conclude that using these models can lead to performance improvements in the risk management of international SME portfolios. Finally, we perform a variety of tests and show that our results remain robust to different distress definitions, estimation techniques and time periods.

Chapter 3 sheds more light on the recent contradictory literature that explores the relationship between default risk and stock returns. We follow a simple and intuitive approach to break down physical PDs into systematic and idiosyncratic components, use the VIX index as a measure of aggregate default risk, and provide European evidence to study the default anomaly.

Initially, we sort stocks into quintile portfolios based on their physical PDs and (in line with the literature that documents a default anomaly) we find that the difference in returns between high and low PD stocks is negative, and that the returns almost monotonically decrease as the PD increases. However, a closer look shows that physical PD is usually a poor measure of exposure to aggregate default risk; in accordance with George and Hwang (2010), we find that stocks in the highest PD quintile have relatively low systematic default risk (SDR) exposure. We then sort stocks into quintile portfolios based on their SDR betas instead. As expected, we find a positive and significant relationship between this measure of default risk and returns. In other words, investors do indeed require a premium to hold stocks that are riskier when aggregate default risk is higher. Interestingly, there are non-monotonic patterns across the SDR beta portfolios. On average, the firms in the low and high SDR beta portfolios are smaller, have a higher level of book-to-market (BM), and higher physical PDs than the firms in medium SDR beta portfolios. We find that the SDR betas are negatively related to the idiosyncratic default risk (IDR) exposures (measured by IDR alphas). Therefore it is the idiosyncratic, not the systematic factors that drive the default anomaly. We confirm this conjecture by showing that stocks sorted on IDR alphas have on average lower returns. Investors do not require compensation to hold stocks with high firm-specific risk because these stocks are a source of portfolio risk diversification. Further analysis with double-sorted portfolios helps us confirm these statements.

Our results suggest that riskier stocks, as measured by the physical PDs, will tend to underperform because they have, on average, lower exposure to aggregate default risk. Their riskiness is mostly idiosyncratic and can be diversified away, thus providing an explanation for the default anomaly typically found in the literature. On the contrary, it is the systematic component of default risk (measured by the SDR betas) that requires a return premium.

Chapter 4 studies the role of both buy out (BO) and venture capital (VC) sponsors in a setting of high information asymmetry, such as the IPO market. These professional insiders may be more capable of taking advantage of such asymmetries compared to insiders of stand-alone companies. But, when the market eventually becomes aware there is then increased risk to reputational capital. BO and VC sponsors may also behave differently from each other. Thus, I differentiate my analysis for each type of PE sponsor and compare BO and VC-backed IPOs with IPOs of stand-alone companies in a matching framework.

I do not find significant differences between these IPOs and matched IPOs of stand-alone companies. The financial situation of both BO and VC-backed companies in the pre-IPO year, as measured by their default risk, is similar to that of their peers. Moreover, PE sponsors do not target their IPOs in “hot” periods more than would managers of stand-alone companies. They also are not more prone to rush their companies into a premature IPO and do not inflate valuations. Finally, PE-backed companies do not default more often post-IPO. This is evidence that PE sponsors are not more likely to seek to sell firms with poor prospects ("unload lemons" in the IPO market.

Chapter 4 provides evidence against media criticism of PE sponsors (e.g. "Rush to get to the front of the IPO queue", Financial Times, 18 February 2014). It can also have important policy implications for the regulatory framework related to PE, such as the Dodd-Frank Act (signed by US President Obama in July 2010). Finally, this paper comes as a timely contribution given the increasing importance of PE-backed IPOs in the market ("Private equity-backed IPOs could hit seven-year high", Financial Times, 29 September 2014).

Chapter 2 Forecasting Distress in European SME Portfolios

2.1 Introduction

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

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

Our paper contributes to the overall literature on corporate credit risk, and on SME risk in particular. It is well known that, unlike larger corporations with easier access to capital markets, SMEs face more challenges in their credit risk modeling. In fact, widely used structural market-based models, such as the distance-to-default (*DD*) measure inspired by Merton (1974), cannot be applied in the non-listed SME setting due to the unavailability of market data. Instead, empirical predictive models such as credit scoring approaches (i.e. Altman, 1968; Edminster, 1972) are most commonly use. Many authors, such as Dietsch and Petey (2004), Berger and Udell (2006), and Beck et al. (2008), note the need for SME specific research. In line with their concerns, Altman and Sabato (2007) (in an early SME study) develop a one-year default prediction model using only accounting information. They apply panel logit estimation on a sample of around 2,000 US SMEs over the period 1994-2002. They find that their model outperforms generic corporate models such as Altman's Z"-score (Altman and Hotchkiss, 2005).

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

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

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

In our study, we find that, in addition to indicators of profitability, coverage, leverage and cash flow, the location and the number of shareholders are important distress determinants for SMEs. We also confirm that systematic factors significantly affect average distress rates in the European economy, a finding that is well-documented in previous US and international literature (Duffie, 2005; Carling et al., 2007; Altman et al., 2012; Jacobson et al., 2013; Laekholm-Jensen et al., 2013). Nevertheless, industry effects often do not demonstrate significance. Moreover, we examine interaction effects between SMEs' size and systematic variables. We find that as SMEs become larger, they are less vulnerable to systematic factors, a finding that is particularly important in light of the current Basel regulations.

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

The paper is organized as follows: Section 2.2 describes the methodology and the reasons for its selection. Section 2.3 describes the dataset, discusses the choice of variables and presents summary statistics. Section 2.4 presents the models and discusses the results, Section 2.5 presents the robustness tests, and Section 2.6 concludes.

2.2 The Methodology

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

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

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

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

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

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

Finally, we account for the survivorship bias, which is the risk that SMEs are more likely to be in our sample if they are survivors and consequently, have lower distress probabilities. Particularly in 2000 (, which is the first year of our sample period), all firms that are present in the database are survivors. This happens because 2000 is the year that our database becomes more complete. As firms enter the database later on, they are always survivors in the first year of their existence in the sample (firms that fail quickly simply are never included in the sample). Thus, we follow a technique similar to Carling et al. (2007) and introduce one more factor, the "duration" variable that accounts for the "time-at-risk" of firms only during the sample period, (i.e. the number of years that a firm stays in the sample). The value of this variable is given by the formula duration=t and is measured in discrete time units. (iI.e., if an SME appears in the sample for three years in total, the value of this variable in the first year is one, in the second year two and in the third year three). By censoring the number of years that a firm existed before it joined the sample, we weight all firms on equal terms and account for duration dependence. This is because, since we allow the time a firm remains in the sample to directly affect the probability of distress, over and above its accounting data and the systematic factors.[1]

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

2.3 The Data

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

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

2.3.1 Definition of Distress

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

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

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

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

It is important to note that the negative equity condition is not used for any of the firm-years of case 1. Now we explain the reason why we add this condition. Our intention in this study is to proxy for distress and not only failure. Thus, we are not only interested in incidents that are strictly determined by legal insolvency procedures. The extended indicator that we use is more appropriate for SMEs because these companies often do not follow such legal procedures at all. A characteristic example is Italy, where there is no clear framework for SMEs to file for insolvency. Even in cases where there is such a framework, filings are not mandatory or they

---

[2] Although we have financial statements data for 2009 and 2010, we do not use them to predict distress for 2010 and 2011 because we do not know which SMEs become distressed during these years. The negative equity condition does not help in this case, since the last available year of our sample is 2010 and we do not know which SMEs disappear the following year, thus we cannot construct our distress indicator for 2010. In Section 2.4, we calculate average distress rates for 2010 and 2011 based on coefficient estimates of the previous years and in Figure 2.1, we plot them together with realized distress rates obtained from an external source.

take a long period of time (e.g. Gilson and Vetsuypens (1993) show in a US study that many filings are missing for bankrupt firms). When these procedures are mandatory, legal insolvency is often related to negative equity. For example, in Germany, firms are obliged to file for bankruptcy once their equity turns negative (Davydenko and Frank, 2008) and in France even earlier, when their equity drops below a certain threshold (LaPorta et al., 1998).

As a result, the proper tracking of the status of SMEs and their distress rates is a very challenging task. There are many different reasons for which an SME can go out of business but owners rarely report these reasons and authorities rarely document them. Watson and Everett (1996) find that small businesses often close for reasons other than distress. For example, a small business can be successful but the owner may still close it voluntarily to accept employment with another company or retire. Headd (2003) finds that only one out of three of start-ups close under conditions that the owners consider unsuccessful. The Amadeus and Orbis databases cooperate in different countries with credit bureaus which provide firm status information. In around 40% of cases though, a firm disappears from the database but the status information remains outdated. In order to separate the cases of closure from the ones of distress for these firms, we need to make a reasonable assumption. This is why we add the negative equity condition.

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

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

To verify our point, in Section 2.5 we perform a robustness test where we exclude all firms that disappear from the sample before 2010 without updated status situation. These are the firms that, under our main distress definition described above, are classified as distressed if their equity is negative in the last year. By excluding these firms, the distress definition in the resulting restricted sample is strictly linked to a legal insolvency procedure. Despite the fact that our sample size decreases significantly, our results remain robust to this alternative distress definition. In Section 2.5 and Appendix 2.3, we report the estimation results as well as comparative statistics between the two distress definitions and samples used. Finally, in unreported results, we replace the negative equity condition with one for negative earnings before interest, taxes, depreciation and amortization (EBITDA). EBITDA is often used in the academic literature as a proxy for operating cashflow. In our sample, 67% of SMEs that have negative equity also report negative EBITDA in the same year. Our results also remain substantially similar under this distress definition.

2.3.2 Sample Selection

In our sample SMEs come from eight European countries, namely Czech Republic, France, Germany, Italy, Poland, Portugal, Spain and the United Kingdom. We select these countries for two reasons: (i) our version of Amadeus and Orbis Europe databases has only European coverage but data are not of the same quality for all countries. (For Scandinavian countries there are very few distress cases for estimation purposes and for most eastern European countries there are very few firms with complete information); (ii) this particular set of countries creates a combination that reflects the variability of SMEs across the EU. This is obvious from Table 2.1, which provides an overview of the key indicators for SMEs in the EU27 and in the countries of our sample. In Italy, Portugal and Spain, SMEs account for larger than EU-average shares of total employment and value added, and present in higher density. This suggests that SMEs in these economies have a greater role than in most EU countries. On the other hand, for France, Germany and the UK, these figures are consistently lower than the EU average. For the Czech Republic and Poland, the share of employment and value added for SMEs is similar.

To study whether the distress determinants differ across Europe and in order to perform comparisons, we split our sample in regional sub-samples. We select the groups based on the following criteria: (ii) the importance of SMEs in the local economies, reflected in Table 2.1; (ii) geography, i.e. west, south, east; (iii) the similarity of the macroeconomic environment, i.e. correlations of macroeconomic variables, currency etc; and lastly, (iv) previous literature. Thus, we form three groups. Group 1 includes the relatively stronger economies of western Europe, namely France, Germany and the UK, group 2 includes the peripheral economies of southern Europe, namely Italy, Portugal and Spain, and group 3 includes two economies from eastern Europe, namely the Czech Republic and Poland. We discuss criterion (i) above and criterion (ii) is clear. Concerning criterion (iii), when we calculate correlations of macroeconomic variables between all country combinations, we find a clear division along the regions. Finally, on criterion (iv), these countries are often bundled together in existing studies (Jaumotte and Sodsriwiboon, 2010; Grammatikos and Vermeulen, 2012; Perego and Vermeulen, 2013).

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

We need criteria (ii)-(iv) to ensure that the companies are independent.[3] Specifically, since we cannot track the subsidiaries and check if the companies still satisfy the criteria to be classified as SMEs once they become subsidiaries, we need to exclude companies that have subsidiaries. Concerning criterion (iv), since the average number of shareholders in our sample is two, we exclude companies with more than ten shareholders as they are possibly outliers. As to criterion (v), we keep companies with at least two years of data in order to be able to lag variables, calculate growth ratios and study the evolution of distress risk. Finally, on criterion (vi), we follow Shumway (2001) and other authors and exclude financial firms from the sample

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

(NACE[4] rev.2 codes from 64 to 68) due to financial firms having reporting practices that preclude combining them with other firms in models using financial information.

After the initial extraction, we apply standard filtering and data cleaning techniques. We first check if missing values can be deduced from other items (i.e. if total assets are missing but fixed and current assets are available, we simply replace total assets with their sum). If the above method does not work, we exclude companies with missing values. We also exclude companies with errors in the data entered (i.e. companies that violate accounting identities). These constrains limit our initial dataset by around 25%.

Our estimation sample consists of 2,721,861 firm-years observations (644,234 firms) out of which 49,355 are distressed. We additionally keep a random one-tenth of the firms from each country as a hold-out sample. The hold-out sample consists of 304,037 firm-year observations (71,823 firms) out of which 5,487 are distressed. Table 2.2 summarizes the properties of our distress indicator for the overall sample and for the regional subsamples. As already mentioned, there is a bias due to the fact that in the beginning of the period (2000-2001), most firms in the database are survivors. It is immediately apparent that Eurozone distress rates are relatively high in 2002-2003, are lower in 2004-2006 and are elevated again from 2007 onwards. This evidence is in accordance with the gloomy business climate in the early years of the last decade, which was followed by an impressive boom of the European economy in 2004-2006 and the subsequent slowdown that started in 2007. The figures are somewhat different for group 3, which consists of two non-Eurozone members. This may be attributed to the fact that the credit supply by banks did not shrink in these countries in the years 2002-2003, as it did in most of the Eurozone. The

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

distressed SMEs are 1.81% of all observations in the overall sample. Group 3 has the highest distress rate (2.4% of all firm-years).

2.3.3 Variables Selection

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

2.3.3.1 Idiosyncratic Variables

Concerning the accounting data, we calculate financial ratios from nine categories: liquidity, profitability, interest coverage, leverage, activity, cash flow, growth (i.e. in sales or profits), asset utilization and employee efficiency.[5] We choose the ratios mainly based on economic intuition and suggestions from past literature. A list of the ratios examined is available upon request. As economic intuition suggests, we expect the probability of distress to be positively related to leverage and negatively related to all other ratio categories.

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

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

and the top one percent to the ninety-ninth percentile, a popular technique known as winsorizing. Finally, because annual reports for SMEs become available with a significant time delay, we lag all ratios in the estimations by 12 months. This means that we assume that data for year $t-1$ become available at the end of year t.

After we calculate the candidate ratios, we follow a standard three-step procedure to select the best for our models. First, we follow Altman and Sabato's (2007) approach and apply the area under the curve (AUC) criterion. The AUC is constructed from the estimated distress probabilities versus the actual status of the firms in each year for all possible cut-off probability values. Thus, we find the AUC for each ratio, applying univariate analysis and keeping those with an AUC above 0.65. Second, we perform correlation analysis to avoid multi-collinearity problems. When the correlation between two ratios is above 0.6, we keep the ratio with the highest AUC. If the difference in the AUC is small, we keep the ratio that was found to be significant in previous studies. Finally, we apply a forward stepwise selection procedure of the remaining ratios, setting the significance level at 10% and performing the likelihood ratio test which is more accurate than the standard Wald test.

Table 2.3 reports summary statistics for the five ratios that are found to be the most effective in predicting distress. A comparison of Panels B and C in Table 2.3 reveals the differences between distressed SMEs. Earnings before taxes to total assets differ substantially across the two groups, suggesting the dominance of unprofitable SMEs in the distressed group. Another striking difference is that the distressed firm-years have, on average, around 130 times lower interest coverage compared to healthy firm-years. Short-term borrowing is also much higher in the case of distressed SMEs. Similarly, turnover to total liabilities ratio is around 180% higher in the healthy firm-years. Finally, the gap between distressed and healthy firm-years in the cash flow ratios indicates the importance of having high cash flows relative to current liabilities.

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

We also account for size, industry type, number of shareholders, location, legal form and age. The European Commission classifies SMEs into three groups based on their number of employees and turnover or total assets: medium-sized enterprises, small enterprises, and micro enterprises. As indicated in Panel A of Table 2.4, our sample is dominated by micro enterprises. In the sixth column of Panel A, the relationship between size and distress risk appears to be non-monotonic, with distress risk relatively stable for medium and small companies and higher for micro companies. This finding is consistent with other studies such as Dietsch and Petey (2004) and is also in line with the argument that smaller companies are more vulnerable to economic fluctuations. To test these predictions, we follow Altman et al. (2010) and employ the natural logarithm of total assets as a proxy for firm size. We also test for other specifications of size, such as total turnover and the number of employees. Additionally, we examine interaction effects between size and the systematic variables that we introduce in the next subsection. For this purpose we use three size dummies (medium, small, micro) and combine them with the systematic variables to test the impact of the macroeconomy on different size groups.

We also control for industry conditions using sector dummies to catch concentration effects. To construct our dummies we use the NACE codes which group industries into 21 major sectors. For estimation purposes though, this classification is too fine. The difficulty here relates to the grouping of sectors into wide sector classes in order to achieve an appropriate degree of homogeneity. It is true that such groupings can always be subject to a certain degree of arbitrariness. In our case, we follow an approach similar to Chava and Jarrow (2004) and form six wide sectors: (i) Sector 1: Agriculture, Mining and Manufacturing, (ii) Sector 2: Transportation, Communication and Utilities, (iii) Sector 3: Construction, (iv) Sector 4: Trade, (v) Sector 5: Accommodation and Food, and (vi) Sector 6: Other services. We select these wide sectors based on different regulatory environments, competition levels and product structures. We also test for alternative groupings but mostly we get insignificant results for more detailed industry classifications. Panel B of Table 2.4 shows the way these broad sectors are partitioned. This initial evidence shows that Accommodation and Food has the highest distress rate and Transportation, Communication and Utilities the lowest.

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

Lastly we examine the age of a smaller sample of firms for which we have the date of establishment. Hudson (1987) finds that companies which are less than ten years old make up the highest proportion of distressed firms. In our sample, the average age at the time of distress is 11.9 years, whereas the average age for healthy firm-years is 15 years. Thus, we expect age to be negatively (but not monotonically) related to distress.

2.3.3.2 Systematic Variables

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

In our models, we use country-specific values and examine systematic variables from three categories: business cycle, credit conditions, and insolvency codes. In Appendix 2.1 we present the variables examined, their expected signs, calculation methods and number of lags, when applied.

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

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

Finally, at this point, we need to elaborate on the effect of bankruptcy laws on distress risk. Davydenko and Franks (2008) examine defaults in three European countries and find differences in insolvency codes among these countries to be important determinants of default outcomes. The World Bank measures the efficiency of insolvency codes in different countries based on the achieved recovery rate, which is the average percentage that claimants recover from an insolvent firm. The recovery rate depends on many factors, such as the time it takes to resolve insolvency proceedings, costs and the outcome of the process. In general, fast, low-cost proceedings and stronger creditor rights characterize the economies with high recovery rates. On the contrary, the more years to resolve an insolvency case, the less friendly the code is and the less likely for the firm to survive during this process. This is also obvious in Appendix 2.2. Countries where the insolvency procedure takes longer (such as the Czech Republic and Poland) score very low as regards the percentage of recoveries. The opposite is true for countries with swift procedures, such as UK and Germany. Thus, we expect distress rates to be negatively related to recovery rates and positively related to the time it takes to resolve insolvency proceedings. The above is also consistent with Acharya et al. (2011), who show that firms in countries with stronger creditor rights (thus higher recoveries) are more conservatively financed (i.e. have less debt).

In order to find among the systematic variables, those which significantly influence the probability of distress for SMEs, we follow a standard procedure. First, we fit the models using only accounting information. Then, we run models that include the ratios and only one systematic variable at a time. We calculate the AUC for each of these models for the overall sample and for the sub-samples, and keep the systematic variables that result to models with the highest AUCs. At this point, we need to account for correlation between the systematic variables. Correlations in this kind of variable are often high, lead to unreasonable signs of the estimated coefficients, and create large changes in the values of these coefficients in response to small changes in the models' specifications. For this reason, between two systematic variables that have a correlation higher than 0.6, we keep the one that results in the model with the highest AUC.

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

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

2.4 The Results

In this section, we present results of models fitted and estimated using the overall sample, models fitted and estimated using the regional sub-samples, and models fitted using the overall sample and estimated using the regional sub-samples. We refer to the models fitted using the overall sample as generic models, and to those fitted using the regional sub-samples as regional models. We identify interesting differences among the European regions, and we show that regional distress models always perform better compared to a generic model estimated using the regional sub-samples.

2.4.1 Generic Models Estimated for the Overall Sample

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

Panel A of Table 2.5 presents the estimated coefficients and chi-square values for the five alternative model specifications. In model I, all firm-specific variables are significant and have the expected signs. Specifically, the probability of distress is negatively related to profitability (earnings before taxes to total assets), coverage (EBITDA to interest expenses), cash flow (cash flow to current liabilities) and activity (turnover to total liabilities) and is positively related to leverage (current liabilities to total assets). Surprisingly, we do not find liquidity ratios as significant in the models. An explanation is that information contained in these ratios is proxied by others. That is, the significance of current liabilities to total assets may indicate that SMEs rely more on short-term borrowing than cash holdings to finance their operational needs. The probability of distress is a decreasing function of the firm size (natural logarithm of total assets), indicating that as the firms become larger, they are less likely to undergo distress (see also Carling et al., 2007). In unreported results, we also test for the non-linear effects of size, by introducing the natural logarithm of squared total assets. We find a positive coefficient, indicating that for the largest SMEs distress risk starts to increase, probably because these companies are more likely to be pursued in liquidation process by their creditors. Two additional interesting findings in accordance with our predictions are that SMEs with less than three shareholders and SMEs in urban areas on average face higher risks. It seems that SMEs with more shareholders receive higher capital support in difficult times. This effect dominates the higher administrative costs that the existence of more shareholders may entail. A possible explanation for the higher risks faced by SMEs in urban areas is that these companies face higher competition (due to geographical proximity) and pay higher rent than their counterparts in non-urban areas. Another reason may be that owners of urban SMEs are less willing to support their enterprises in times of difficulties. This strategic distress caused because it is a more often a viable option for business owners to close the business and find employment elsewhere. [6] These effects seem to out-weigh the fact that there is a larger customer base available for urban SMEs.

In model II, the firm-specific variables retain their significance and signs once the systematic variables are added. We identify five systematic variables as doing the best overall job in predicting distress, namely the FX rate change, the unemployment, the economic sentiment indicator, and the change in bank lending. As we hypothesized, an appreciation of the currency, an increase in the economic sentiment indicator and an increase in lending by banks result in lower distress rates. Conversely, an increase in unemployment and a greater number of years to achieve insolvency resolution result in higher distress rates. To assess the usefulness of the systematic variables, we perform a likelihood ratio test for the nested models I and II. The null hypothesis that the coefficients of these variables are jointly equal to zero is strongly rejected, as indicated in Table 2.5.

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

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

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

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

In model V, we introduce firm age and test its effect on distress probability for the slightly smaller sample for which we have available data on age. We find, in accordance with previous literature, that older firms are safer. We also follow Altman et al. (2010) and check for non-linear effects of age. As in their study, we find a positive and statistically significant coefficient for a dummy variable equal to one if SMEs are between three and nine years old.

We notice that the pseudo-R^2 (McFadden's R^2) is increasing along the different model specifications, indicating a better fit as we add more variables. The pseudo-R^2 values may look low when compared to R^2 values of linear regression models, but such low values are normal in logistic regression (Hosmer and Lemeshow; 2000). In order to evaluate more closely the performance of our models, we perform in-sample and out-of-sample testing. We employ two widely used measures, the Hosmer and Lemeshow grouping based on estimated distress probabilities and the area under the curve (AUC).

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

The AUC is constructed from the estimated distress probabilities versus the actual status of the firms in each year for all possible cut-off probability values. Specifically, the curve plots the ratio of correctly classified distressed firms to actual distressed firms (sensitivity) and the ratio of wrongly classified healthy firms to actual healthy firms (1 - specificity) for all possible cut-offs. The AUC ranges from zero to one. A model with an AUC close to 0.5 is considered a random model with no discriminatory power. An AUC of 0.7 to 0.8 represents good discriminatory power, an AUC of 0.8 to 0.9 very good discriminatory power and an AUC over 0.9 is exceptional and extremely unusual. The AUC criterion is an improvement to the traditional classification tables that rely on a single cut-off point to classify distressed and healthy firms. Several statistics are equivalent to the AUC, such as the accuracy ratio, the Gini coefficient and the Mann-Whitney-Wilcoxon test (Engelmann et al., 2003). While the Hosmer and Lemeshow method assesses mainly calibration, the AUC assesses discrimination.

Panel B of Table 2.5 presents the results of the in-sample tests. According to the Hosmer - Lemeshow grouping, the percentage of distressed firms in the last three deciles increases from model I to model II (75.83% to 76.59%). Also, the percentage of distressed firms in the first five deciles drops (11.38% to 11.09%). These show that adding the systematic variables improves performance both in terms of an increase in the correct classification of distressed firms and a decrease in the incorrect classification of healthy firms. AUC also increases from 0.824 to 0.838. This result is stronger than those achieved by previous studies in the literature. Specifically in Altman et al. (2010) this figure ranges between 0.78 and 0.80. When it comes to model III, it only modestly outperforms model II. Specifically, by taking industry effects into account, the AUC remains almost the same and the percentage of distressed firms in the last three deciles increases slightly (76.59% to 76.66%). Given these results, controlling for industry effects improves performance only marginally, once we have already accounted for systematic factors. When we add interaction effects between size and systematic factors, we notice a further increase in the percentage of distressed firms in the last three deciles (76.66% to 77.06%). AUC also increases from 0.839 to 0.843. Moving to model V, it seems that age also helps slightly. However, we cannot directly compare model IV to model V since model V is estimated with a smaller sample.

Panel C of Table 2.5 presents the results of the out-of-sample tests. Out-of-sample testing is challenging since improvements in the in-sample fit can be a result of over-fitting of the original data. As seen, all results follow the same patterns as for the in-sample tests.

The in-sample and out-of-sample tests provide evidence that distress is captured more successfully with systematic variables and their interaction effects than with industry effects (which help only marginally). In unreported results, we also run a model where we include only firm-specific information (model I) and the industry dummies. As expected, this model performs worse than model II, which includes firm-specific information and the systematic factors.

The importance of systematic variables in distress prediction is also demonstrated in Figure 2.1, where we plot the predicted and observed distress rate for the period 2000 to 2011. The predicted distress rate is the simple average of the probabilities of distress of all firms in each period. Since we have financial but not distress information for 2010 and for only a few companies in 2011, we do the following: (i) we use the estimated coefficients from 2000-2009 to predict the distress rate for 2000-2011; (ii) we use the in-sample observed distress rate for 2000-2009 and we obtain the observed distress rate for 2010-2011 from Creditreform, the largest private credit bureau in Germany that gathers statistics on insolvencies in Europe. Thus years 2010 and 2011 provide out-of-sample evidence. The columns denote recession periods in the Eurozone as indicated by the OECD. The graph shows that in model I, where only firm-specific variables are included, the predicted distress rates follow a smooth upward trend, but do not co-vary with the observed distress rates. It is the systematic variables (present in models II, III and IV) that shift the mean of the distress distribution and are able to capture distress-clustering during recessions. When systematic variables are included, predicted distress rates move together with observed ones and vary greatly with the business cycle, increasing with downturns and lowering with upturns. Once again, industry effects do not seem to provide additional improvements. These findings are in accordance with previous literature (Carling et al., 2007; Jacobson et al., 2005, 2013; Laekholm-Jensen et al., 2014).

2.4.2 Generic and Regional Models Estimated for the Regional Subsamples

Now, we turn our analysis to the regional sub-samples presented in Section 2.3. First, we use the generic specifications of subsection 2.4.1 and estimate the generic models for the regional subsamples. Our preferred model is model II of Table 2.5, because it considers both idiosyncratic and systematic variables, performs very well (AUC of 0.8382 and 65.44% of distressed SMEs in the last three Hosmer Lemeshow deciles) and, at the same time, has a simple specification. We ignore industry effects (model III) because they do not add much in terms of performance improvements. We also ignore interaction effects (model IV) and age (model V) for reasons of parsimony. Later, we fit idiosyncratic and systematic variables for the regional sub-samples and estimate three regional models. Lastly, we compare the generic models estimated for the regional sub-samples with the regional models.[7]

Table 2.6 presents the results. In accordance with our hypothesis, we document performance improvements when we switch to the regional models. Interestingly, we find that the firm-specific variables identified as the most important in predicting distress in the regional models are exactly the same as in the generic models. This is evidence that SMEs across Europe are sensitive to the same idiosyncratic factors. This does not hold in the case of systematic factors. Specifically, we document regional variations in the vulnerabilities to systematic factors,

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

according to region-specific conditions and characteristics. Moreover, we do not find that the years taken to resolve insolvency variable adds predictive power in the regional models. This can partly be due to regional groups being relatively homogeneous with respect to their insolvency regimes.

For group 1 (France, Germany, U.K.), the models are estimated from a sample of 165,786 SMEs (801,536 firm-year observations), which include 14,177 distressed SMEs. When we move from the generic to the regional model, we document small changes in the coefficient sizes of the idiosyncratic variables. We also find that we need only two (instead of five) systematic variables to predict distress. These variables are the bank lending and the GDP growth criteria. Both bank lending and GDP growth have significant coefficients and are, as expected, negatively related to the distress rate. Lower GDP growth means lower growth in sales by firms and thus an increased distress probability. Interestingly, even with less systematic variables, the regional model achieves higher performance than the generic one. Specifically, the percentage of distressed firms in the last three deciles increases from the generic to the regional model (72.94% to 74.16%) and the AUC increases (0.806 to 0.825). Out-of-sample performance improvements are similar as in the case of in-sample results. The above provides evidence that SMEs in the countries of group 1, which consists of some of the strongest EU economies, are less sensitive to the macroeconomic situation. This is in accordance with the finding in subsection 2.4.1, that large SMEs are less vulnerable to the macro-economic situation compared to small SMEs, since SMEs in group 1 are, on average, larger.

For group 2 (Italy, Portugal, Spain), the models are estimated from a sample of 429,978 SMEs (1,741,707 firm-year observations), which include 30,900 distressed SMEs. When we move from the generic to the regional model, we identify almost the same systematic factors as being the most useful for predicting distress. The years to resolve insolvency variable is replaced by the balance of payments variable, since these countries often suffer from current account deficits. It is interesting to note that, in accordance with our predictions, group 2 is vulnerable to more macroeconomic factors compared to group 1. The reason for this can be the generally less favourable economic climate in the economies of group 2 during the years of this study. Here, the regional and generic models have a very similar performance. Specifically, the regional model only modestly outperforms the generic one (a 0.05% improvement in the correct classification for distressed firms and a 0.07 improvement in AUC). This happens because group 2 represents 64% of the overall sample, thus, it mainly drives the results of the generic model.

For group 3 (Czech Republic, Poland), the models are estimated from a sample of 48,470 SMEs (178,618 firm-year observations), which include 4,278 distressed SMEs. When we move from the generic to the regional model, coefficient sizes of the idiosyncratic variables differ slightly and we observe an interesting new set of systematic variables. We find the FX volatility, the 10-year government bond yield and the GDP growth variables as the most useful systematic variables in predicting distress. With respect to the volatility of the exchange rate, higher volatility is positively related to distress (see also Nam et al., 2008). Interestingly, as we hypothesized, it seems that, for the non-Eurozone countries of group 3, the stability of their national currencies plays a crucial role in the solvency of SMEs. This is presumably due to the fact that a very volatile FX rate in these economies increases instability and thus creates uncertainty about future economic conditions. Concerning the 10-year government bond yield variable, it enters with a positive coefficient. Thus, a higher interest rate is positively related to distress. Government bond yields are systematically higher in the countries of group 3 compared to the rest of the sample for the years of the study, indicating the higher sovereign risk (country premium) for these economies. As before, GDP growth is negatively related to distress. According to the Hosmer-Lemeshow grouping, the percentage of distressed firms in the last three deciles increases from the generic to the regional model (81.79% to 82.54%). AUC also increases (0.853 to 0.875). Clearly, the specific set of systematic variables helps in better capturing distress risk. The out-of-sample results give the same picture.

The above evidence shows that the fit is improved when we change some of the macroeconomic co-variates as we move from the generic to the regional models. This indicates that regional models are better able to capture the systematic effects. Although the improvements might seem small, Figures 2.2 and 2.3 give a better sense of the comparative performance.

Figure 2.2 plots the predicted distress rate based on the regional and generic models of Table 2.6, along with the observed distress rate for each group. It is obvious that the predicted distress rates from the three regional models match better the observed values, compared to the predicted distress rate from the generic models. For group 1, the generic model underestimates the distress rate for the early years of our study (before 2004) and overestimates it later on (from 2008 onwards). In the case of group 2, the two time-series are very similar. This is probably due to the vast majority of companies in our sample belonging to group 2, thus the regional and generic model for this group include almost the same co-variates and give almost identical predictions. Finally, for group 3, the generic model overestimates the distress rates for the years 2003-2004 and co-moves with the regional model (and the observed values) in later years. The years 2010-2011 provide out-of-sample evidence for group 1 and 2. The predicted distress rate closely follows the pattern of the observed one, specifically the falling trend in 2010-2011. Please note that for group 3, we lack distress information for 2010-2011, thus the extension is not possible.

Figure 2.3 plots the aggregate time series for four macroeconomic variables. They are economic sentiment, unemployment (in percentage terms), the balance of payments (as a percentage of the GDP), and foreign exchange rate volatility. The economic sentiment, the unemployment and the balance of payments variables are average values for the countries in each group. The foreign exchange rate is in relation to the US dollar. We report volatility for each currency separately (euro, British pound, Czech koruna and Polish zloty).

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

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

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

Finally, the volatility of foreign exchange rates against the dollar follow similar trends for all currencies, but it is the Czech koruna and Polish zloty that have the highest volatility values.

Thus, in the regional model for group 3, we identify this variable as a significant determinant of the SME distress rate.

2.5 Robustness Tests

2.5.1 Definition of Distress

In Section 2.3, we discuss that in around 40% of our sample (254,887 out of 644,234 firms), a firm disappears from the database before 2010 but the status information remains outdated. We also extensively discuss the challenges in tracking properly the status of SMEs that lead us to adopt an assumption for this 40% of firms. Thus, under our main distress definition, in order to separate the cases of closure from the ones of distress, we assume that the last available firm-year of these firms is distressed if equity is negative in this last year. Because this assumption influences a large percentage of our sample, the estimation results can be sensitive to it. Therefore, in this section, we perform a robustness test where we exclude the 254,887 firms (1,127,428 firm-years) that disappear from the sample before 2010 without updated status situation. So, under this alternative distress definition, distress is strictly linked to only a legal insolvency procedure. The remaining sample includes 1,594,433 firm-years (389,374 firms) out of which 12,362 are distressed. Appendix 2.3 reports comparative statistics between the two distress definitions and samples used.

Table 2.7 reports the estimation results for all countries and for the regional groups under both definitions for comparison purposes. Despite the fact that our sample size decreases significantly and distress rates are much lower, our results remain robust. Almost all variables retain their signs and significance. In a few cases that the sign flips, coefficients do not demonstrate significance. The exception is size, which has a significantly negative coefficient under the main distress definition and a significantly positive coefficient for the overall sample (as well as groups 1 and 2) under the alternative distress definition. An explanation for this can be its non-linear effect. Specifically, in page 23, we mention that we find a positive coefficient for squared size, indicating that for very large SMEs, distress risk starts to increase, probably because these companies are more likely to be pursued by their creditors in liquidation procedures. Further supporting this argument, we find that the 12,362 distressed cases under the alternative distress definition come from 200% larger companies than the 49,355 distressed cases under the main definition.

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

2.5.2 Estimation Technique

In addition to the multi-period logit model developed by Shumway (2001), we apply the Cox proportional hazard model (Cox, 1975) that makes different assumptions about the hazard function. A hazard model is a type of survival model, in which the co-variates are related to the time that passes before some event occurs (in this case distress). Specifically, we follow Laerkholm-Jensen et al. (2013) and estimate a fully parametric model with a constant baseline intensity, since the usual Cox semi-parametric model does not allow us to simultaneously identify the vector of macroeconomic coefficients as well as the time-varying baseline intensity.

Table 2.8 reports the estimation results for the overall sample and for the regional sub-samples under both techniques for comparison purposes. All our results remain robust when we instead apply the Cox model. Specifically, all regression coefficients retain their sign and significance. The sizes of the coefficients are very similar as well.

2.5.3 Time Periods

We also split our sample into four rolling window periods (each one lasting five years). We find that whereas sensitivities to idiosyncratic factors remain relatively stable over time, coefficients of systematic variables are more volatile, responding to changes in the prevailing macroeconomic conditions.

In this section, we estimate the generic model for the overall sample over four rolling windows, each five years long during the period 2002-2009. We perform this analysis for two reasons: first, in order to examine the stability of coefficients through time; and secondly, to further test performance. This time, we evaluate predictive power over exactly the next year following each model's estimation period as well as over the last year of our sample (2009).

Panel A of Table 2.9 presents the estimation results of the four rolling windows over the period 2002-2009, as well as of the overall sample (period 2000-2009) for comparison purposes. Coefficients of firm-specific variables are always significant and keep the same signs along the different windows, but there is relative variation in their magnitudes. The only puzzling result is the positive coefficient of size in the 2004-2008 window, but it seems that this result is sample specific. Coefficients of systematic variables follow the same patterns but display higher volatility, presumably as a result of the changing economic conditions during the period of the study. The years to resolve insolvency are negatively related to distress in the 2002-2006 window but this is probably also sample specific since distress rates are increasing quite strongly from 2002 to 2003 (Table 2.2) but insolvency regimes remain stable or improve.

Panels B and C of Table 2.9 present the out-of-the-sample performance of the estimated rolling windows. Specifically, Panel B presents performance over the next year following the estimation period and Panel C presents performance over the last sample year (2009). In Panel

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

2.6 Concluding Remarks

The paper explores the performance of distress prediction hazard models for non-financial SMEs using a dataset from eight European countries over the ten-year period 2000-2009. We find that (in addition to financial indicators whose importance is noted in past studies) the location and number of shareholders are important determinants of SMEs’ distress probabilities. We validate the superiority of models that incorporate macroeconomic dependencies, suggested by previous research, also in the case of SMEs but do not find strong evidence that industry effects significantly improve prediction accuracy. We also examine interaction effects between SMEs’ size and systematic variables. We find that as SMEs become larger, they are less vulnerable to the macroeconomic situation, contrary to what Basel regulations assume. Interestingly, when we split our sample in regional sub-samples, we show that SMEs across Europe are sensitive to the same firm-specific factors, but we identify significant regional variations in the selection and importance of macro variables. Specifically, macro variables differ among European regions based on region-specific conditions and characteristics. Since our regional distress models always perform better than a generic model estimated for the regional sub-samples, we conclude that their use can lead to performance improvements in the risk management of international SME portfolios. Finally, we perform a variety of tests and show that our results remain robust to different distress definitions, estimation techniques and time periods.

2.7 Tables of Chapter 2

Table 2.1

Key Indicators

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

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

Table 2.2

Distressed SMEs as Percentage of Total SMEs

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

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

Table 2.3
Summary Statistics

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

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

Table 2.4
SMEs by Size and Industry

Panel A. Size classification

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

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

Panel B. Industry classification (wide sectors)

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

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

Table 2.5

Generic Models Estimated for the Overall Sample

Panel A. Estimation results

The models are estimated for 2000-2009 with lagged yearly observations using the multi-period logit technique. The data-set includes non-financial SMEs from eight European economies. Parameter estimates are given first followed by chi-square values in parentheses. There are 644,234 firms in the sample (2,721,861 firm-year observations) out of which 49,355 distressed.

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

Table 2.5. Cont.

	Model I	Model II	Model III	Model IV	Model V
Firm-year observations	2,721,861	2,721,861	2,721,861	2,721,861	2,652,157
Firms	644,234	644,234	644,234	644,234	620,872
Distressed firms	49,355	49,355	49,355	49,355	47,841
Pseudo R-squared	0.147	0.171	0.171	0.178	0.187
Log likelihood	-210,601.30	-204,638.50	-204,538.30	202,880.11	194,837.44
Wald test	78,110.8***	84,259.5***	84,526.8***	85,305.9	81,789.3***
Likelihood ratio test		11,925.57***	200.45***	3,316.36	16,085.34***
* p<0.05, ** p<0.01, *** p<0.001					
Panel B. In-sample prediction tests					
Hosmer-Lemeshow decile					
1 to 5	11.38%	11.09%	10.96%	10.67%	10.24%
8	11.20%	11.16%	11.25%	10.91%	10.33%
9	17.46%	17.86%	17.84%	17.83%	17.34%
10	47.17%	47.58%	47.57%	48.32%	49.49%
8 to 10	**75.83%**	**76.59%**	**76.66%**	**77.06%**	**77.16%**
Area under the ROC curve	0.824	0.838	0.839	0.843	0.857
Panel C. Out-of-sample prediction tests					
A hold-out sample of 71,823 European SMEs (304,037 firm-year observations) is used.					
Hosmer-Lemeshow decile					
1 to 5	11.46%	11.35%	11.26%	10.30%	10.15%
8	11.24%	10.53%	10.66%	11.45%	10.71%
9	17.51%	18.70%	18.53%	18.15%	17.16%
10	46.78%	47.04%	47.04%	47.95%	48.95%
8 to 10	**75.54%**	**76.27%**	**76.23%**	**77.55%**	**76.82%**
Area under the ROC curve	0.823	0.837	0.837	0.844	0.847

Table 2.6

Generic and Regional Models Estimated for the Regional Subsamples

Panel A. Estimation results

The models are estimated for 2000-2009 data with lagged yearly observations using the multi-period logit technique. The data-set is limited to non-financial SMEs. Parameter estimates are given first followed by chi-square values in parentheses. Group 1 has 165,786 French, German and British SMEs (801,536 firm-year observations) out of which 14,177 distressed. Group 2 has 429,978 Italian, Portuguese and Spanish SMEs (1,741,707 firm-year observations) out of which 30,900 distressed. Group 3 has 48,470 Czech and Polish SMEs (178,618 firm-year observations) out of which 4,278 distressed.

	Group 1				Group 2				Group 3			
	Generic Model		Regional Model		Generic Model		Regional Model		Generic Model		Regional Model	
Earnings before taxes to total assets	-1.0667***	(-15.97)	-1.077***	(-15.54)	-0.681***	(-10.67)	-0.677***	(-10.61)	-0.534***	(-4.50)	-0.547***	(-4.68)
EBITDA to interest expenses	-0.0000470***	(-11.10)	-0.0000483***	(-11.31)	-0.0000472***	(-9.80)	-0.0000468***	(-9.82)	-0.0000507***	(-4.62)	-0.0000509***	(-4.55)
Current liabilities to total assets	1.908***	(71.61)	1.916***	(72.99)	1.216***	(69.49)	1.217***	(69.66)	1.415***	(33.07)	1.397***	(32.84)
Cash flow to current liabilities	-0.236***	(-4.15)	-0.196**	(-3.16)	-0.654***	(-9.51)	-0.650***	(-9.43)	-0.336***	(-2.75)	-0.314**	(-2.64)
Turnover to total liabilities	-0.106***	(-14.72)	-0.101***	(-14.07)	-0.245***	(-30.10)	-0.249***	(-30.41)	-0.171***	(-13.05)	-0.175***	(-13.28)
Size (ln(totals assets))	-0.0226***	(-3.05)	-0.00559	(-0.75)	-0.101***	(-16.95)	-0.107***	(-17.72)	-0.0587***	(-4.60)	-0.0754***	(-6.04)
Dummy equal to 1 if shareholders are more than 2	-0.0781***	(-3.44)	-0.0812***	(-3.62)	-0.334***	(-20.51)	-0.324***	(-19.96)	-0.344***	(-7.49)	-0.347***	(-7.53)
Dummy equal to 1 if SME is located in urban area	0.151***	(4.82)	0.174***	(5.68)	0.103***	(6.50)	0.101***	(6.43)	0.351***	(9.75)	0.358***	(9.92)
Loans granted to non-financial sector (% change)	-2.388***	(-14.95)	-4.611***	(-25.53)	-0.268	(-1.29)	-3.378***	(-30.14)	-6.203***	(-21.64)		
Years to resolve insolvency proceedings	-1.206***	(-21.27)			1.171***	(18.92)			-0.237***	(-20.25)		
GDP growth (% change)			-5.595***	(-9.44)							-11.62***	(-22.52)
FX rate (% change)	-2,052.2***	(-49.78)			-2,403.5***	(-46.89)	-2276.6***	(-44.99)	340.15***	(3.06)		
Unemployment	19.425***	(13.59)			13.441***	(27.02)	6.176***	(24.91)	-22.213***	(-15.94)		
Economic sentiment	-0.0206***	(-15.16)			-0.0279***	(-23.62)	-0.0256***	(-21.08)	-0.0031	(-0.93)		
Balance of Payments (% GDP)							-42.082***	(-16.28)				
FX rate volatility											122.6***	(12.11)
10-year government bond yield											25.43***	(14.63)
Constant	Yes		Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes		Yes	

Table 2.6. Cont.

	Group 1		Group 2		Group 3	
	Generic Model	Regional Model	Generic Model	Regional Model	Generic Model	Regional Model
Firm-year observations	801,536	801,536	1,741,707	1,741,707	178,618	178,618
Firms	165,786	165,786	429,978	429,978	48,470	48,470
Distressed firms	14,177	14,177	30,900	30,900	4,278	4,278
Pseudo R-squared	0.150	0.150	0.170	0.177	0.214	0.250
Log likelihood	-61,573.50	-60,538.70	-131,451.70	-127,673.50	-15,878.40	-15,147.90
Wald test	19,302.49***	20,225.9***	55,513.51***	55,783.8***	8,099.98***	8,083.9***
* p<0.05, ** p<0.01, *** p<0.001						
Panel B. In-sample prediction tests						
Hosmer-Lemeshow decile						
1 to 5	14.16%	13.85%	9.50%	9.17%	7.25%	7.22%
8	9.85%	9.47%	11.42%	11.86%	8.86%	9.70%
9	14.14%	14.02%	18.90%	18.67%	17.58%	17.16%
10	48.95%	50.67%	47.17%	47.01%	55.35%	55.68%
8 to 10	**72.94%**	**74.16%**	**77.49%**	**77.54%**	**81.79%**	**82.54%**
Area under the ROC curve	0.806	0.825	0.841	0.848	0.853	0.875
Panel C. Out-of-sample prediction tests						
For Group 1, a hold-out sample of 18,449 French, German and British SMEs (88,957 firm-year observations) is used. For Group 2, a hold-out sample of 48,034 Italian, Portuguese and Spanish SMEs (195,236 firm-year observations) is used. For Group 3, a hold-out sample of 5,340 Czech and Polish SMEs (19,844 firm-year observations) is used.						
Hosmer-Lemeshow decile						
1 to 5	13.84%	13.47%	9.23%	9.11%	7.49%	7.73%
8	9.47%	8.79%	11.06%	10.80%	9.37%	8.90%
9	13.22%	15.01%	19.02%	19.57%	16.86%	16.63%
10	47.79%	48.95%	47.66%	47.38%	57.61%	58.08%
8 to 10	**70.48%**	**72.76%**	**77.73%**	**77.75%**	**81.26%**	**83.61%**
Area under the ROC curve	0.805	0.824	0.843	0.850	0.841	0.868

Table 2.7

Robustness Test with Different Distress Definitions

The models are estimated for 2000-2009 data with lagged yearly observations using the multi-period logit technique. The data-set is limited to non-financial SMEs. Parameter estimates are given first followed by chi-square values in parentheses. According to the main distress definition, a firm-year is distressed if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) either appears with one of the statuses defaulted, in receivership, bankrupt, in liquidation or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. In the alternative distress definition, we exclude all firms that disappear from the sample before 2010 without updated status situation. These include firms that, under the main distress definition are classified as distressed if their equity is negative in the last year. Thus, the alternative distress definition is strictly linked to a legal insolvency procedure.

	Overall Sample - Generic Model				Group 1 - Regional Model			
	Main definition		Alternative definition		Main definition		Alternative definition	
Earnings before taxes to total assets	-0.770***	(-15.89)	-0.699***	(-5.43)	-1.077***	(-15.54)	-0.525***	(-3.36)
EBITDA to interest expenses	-0.0000450***	(-14.49)	-0.0000289***	(-6.33)	-0.0000483***	(-11.31)	-0.0000333***	(-6.15)
Current liabilities to total assets	1.420***	(103.81)	1.536***	(49.76)	1.916***	(72.99)	2.007***	(44.26)
Cash flow to current liabilities	-0.485***	(-9.14)	-1.277***	(-11.35)	-0.196**	(-3.16)	-1.149***	(-8.13)
Turnover to total liabilities	-0.177***	(-36.08)	-0.035***	(-5.67)	-0.101***	(-14.07)	-0.017*	(-2.17)
Size (ln(totals assets))	-0.0940***	(-22.95)	0.376***	(45.80)	-0.00559	(-0.75)	0.308***	(27.33)
Dummy equal to 1 if shareholders are more than 2	-0.274***	(-21.99)	-0.214***	(-9.48)	-0.0812***	(-3.62)	-0.0585	(-1.79)
Dummy equal to 1 if SME is located in an urban area	0.141***	(10.85)	0.0341	(1.15)	0.174***	(5.68)	0.239***	(5.05)
FX rate (% change)	-1,686.8***	(-59.04)	-1,357.1***	(-27.83)				
Unemployment	1.883***	(12.39)	1.502***	(28.97)				
Economic sentiment indicator	-0.0259***	(-35.03)	-0.0047***	(-3.22)				
Loans granted to non-financial sector (% change)	-4.414***	(-58.29)	-5.194***	(-34.27)	-4.611***	(-25.53)	-6.404***	(-20.77)
Years to resolve insolvency proceedings	0.0949***	(27.40)	0.732***	(21.49)				
GDP growth (% change)					-5.595***	(-9.44)	0.963	(0.00)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	2,721,861		1,594,433		801,536		332,547	
Firms	644,234		389,347		165,786		66,306	
Distressed firms	49,355		12,362		14,177		5,646	
Pseudo R-squared	0.171		0.115		0.150		0.098	
Log likelihood	-204,638.50		-60,050.11		-60,538.70		-25,683.62	
Wald test	84,259.5***		16,563.99***		20,225.9***		5,359.19***	
* p<0.05, ** p<0.01, *** p<0.001								
Area under the ROC curve	0.838		0.794		0.825		0.776	

Table 2.7. Cont.

	Group 2 - Regional Model				Group 3 - Regional Model			
	Main definition		Alternative definition		Main definition		Alternative definition	
Earnings before taxes to total assets	-0.677***	(-10.61)	-1.248***	(-6.63)	-0.547***	(-4.68)	-3.092***	(-4.45)
EBITDA to interest expenses	-0.0000468***	(-9.82)	-0.0000507***	(-5.21)	-0.0000509***	(-4.55)	0.0000009	(0.03)
Current liabilities to total assets	1.217***	(69.66)	1.126***	(26.39)	1.397***	(32.84)	1.204***	(3.21)
Cash flow to current liabilities	-0.650***	(-9.43)	-1.111***	(-7.27)	-0.314**	(-2.64)	-0.149***	(-0.68)
Turnover to total liabilities	-0.249***	(-30.41)	-0.0925***	(-7.25)	-0.175***	(-13.28)	-0.0195	(-0.86)
Size (ln(totals assets))	-0.107***	(-17.72)	0.462***	(35.44)	-0.0754***	(-6.04)	-0.1228	(-1.16)
Dummy equal to 1 if shareholders are more than 2	-0.324***	(-19.96)	-0.368***	(-11.36)	-0.347***	(-7.53)	-0.229	(-1.15)
Dummy equal to 1 if SME is located in urban area	0.101***	(6.43)	0.0037	(0.10)	0.358***	(9.92)	0.354	(1.20)
Loans granted to non-financial sector (% change)	-3.378***	(-30.14)	-4.482***	(-20.47)				
GDP growth (% change)					-11.62***	(-22.52)	0.700	(0.11)
FX rate (% change)	-2276.6***	(-44.99)	-808.86***	(-11.98)				
Unemployment	6.176***	(24.91)	6.709***	(9.16)				
Economic sentiment	-0.0256***	(-21.08)	-0.01888***	(-6.56)				
Balance of Payments (% GDP)	-42.082***	(-16.28)	-30.575***	(-35.61)				
FX rate volatility					122.6***	(12.11)	308.32***	(4.85)
10-year government bond yield					25.43***	(14.63)	57.14*	(2.85)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	1,741,707		1,185,258		178,618		76,628	
Firms	429,978		302,959		48,470		20,082	
Distressed firms	30,900		6,338		4,278		378	
Pseudo R-squared	0.177		0.127		0.250		0.214	
Log likelihood	-127,673.50		-32,418.66		-15,147.90		-1,247.08	
Wald test	55,783.8***		9,052.84***		8,083.9***		611.51	
* p<0.05, ** p<0.01, *** p<0.001								
Area under the ROC curve	0.848		0.847		0.879		0.816	

Table 2.8

Robustness Test with Different Estimation Techniques

The models are estimated for 2000-2009 data with lagged yearly observations using the multi-period logit and Cox proportional hazard techniques. The data-set is limited to non-financial SMEs. Parameter estimates are given first followed by chi-square values in parentheses. The Cox proportional hazard model makes different assumptions about the hazard function. We follow Laerkholm-Jensen et al. (2013) and estimate a fully parametric model with a constant baseline intensity, since the usual Cox semi-parametric model does not allow us to simultaneously identify the vector of macroeconomic coefficients as well as the time-varying baseline intensity.

	Overall Sample - Generic Model				Group 1 - Regional Model			
	Logit		Cox		Logit		Cox	
Earnings before taxes to total assets	-0.770***	(-15.89)	-0.496***	(-11.30)	-1.077***	(-15.54)	-0.719***	(-10.21)
EBITDA to interest expenses	-0.0000450***	(-14.49)	-0.0000413***	(-13.79)	-0.0000483***	(-11.31)	-0.0000463***	(-11.11)
Current liabilities to total assets	1.420***	(103.81)	1.197***	(94.38)	1.916***	(72.99)	1.653***	(69.97)
Cash flow to current liabilities	-0.485***	(-9.14)	-0.667***	(-13.15)	-0.196**	(-3.16)	-0.367***	(-4.69)
Turnover to total liabilities	-0.177***	(-36.08)	-0.183***	(-37.28)	-0.101***	(-14.07)	-0.088***	(-12.91)
Size (ln(totals assets))	-0.0940***	(-22.95)	-0.084***	(-21.64)	-0.00559	(-0.75)	0.01882**	(2.95)
Dummy equal to 1 if shareholders are more than 2	-0.274***	(-21.99)	-0.239***	(-20.17)	-0.0812***	(-3.62)	-0.0635**	(-2.98)
Dummy equal to 1 if SME is located in an urban area	0.141***	(10.85)	0.096***	(7.95)	0.174***	(5.68)	0.122***	(4.24)
FX rate (% change)	-1,686.8	(-59.04)	-1,445.2***	(-52.53)				
Unemployment	1.883***	(12.39)	3.084***	(21.66)				
Economic sentiment indicator	-0.0259***	(-35.03)	-0.0066***	(-8.46)				
Loans granted to non-financial sector (% change)	-4.414***	(-58.29)	-3.624***	(-53.55)	-4.611***	(-25.53)	-2.713***	(-19.34)
Years to resolve insolvency proceedings	0.0949***	(27.40)	0.0993***	(30.72)				
GDP growth (% change)					-5.595***	(-9.44)	-3.844***	(-7.45)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	2,721,861		2,721,861		801,536		801,536	
Firms	644,234		644,234		165,786		165,786	
Distressed firms	49,355		49,355		14,177		14,177	
Log likelihood	-204,638.50		-91,145.41		-60,538.70		-25,643.69	
Wald test	84,259.5***		97,333.98***		20,225.9***		26,984.78***	

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

Table 2.8. Cont.

	Group 2 - Regional Model				Group 3 - Regional Model			
	Logit		Cox		Logit		Cox	
Earnings before taxes to total assets	-0.677***	(-10.61)	-0.421***	(-7.54)	-0.547***	(-4.68)	-0.304**	(-2.86)
EBITDA to interest expenses	-0.0000468***	(-9.82)	-0.0000413***	(-9.20)	-0.0000509***	(-4.55)	-0.0000451***	(-3.74)
Current liabilities to total assets	1.217***	(69.66)	1.026***	(63.38)	1.397***	(32.84)	1.135***	(27.85)
Cash flow to current liabilities	-0.650***	(-9.43)	-0.803***	(-13.08)	-0.314**	(-2.64)	-0.416**	(-3.26)
Turnover to total liabilities	-0.249***	(-30.41)	-0.239***	(-30.81)	-0.175***	(-13.28)	-0.171***	(-12.25)
Size (ln(totals assets))	-0.107***	(-17.72)	-0.097***	(-17.41)	-0.0754***	(-6.04)	-0.0355**	(-3.17)
Dummy equal to 1 if shareholders are more than 2	-0.324***	(-19.96)	-0.302***	(-19.71)	-0.347***	(-7.53)	-0.283***	(-6.28)
Dummy equal to 1 if SME is located in urban area	0.101***	(6.43)	0.079***	(5.47)	0.358***	(9.92)	0.276***	(8.10)
Loans granted to non-financial sector (% change)	-3.378***	(-30.14)	-4.178***	(-40.2)			-10.25***	(-17.42)
GDP growth (% change)					-11.62***	(-22.52)	-10.25***	(-17.42)
FX rate (% change)	-2276.6***	(-44.99)	-2083.1***	(-42.22)				
Unemployment	6.176***	(24.91)	5.488***	(24.11)				
Economic sentiment	-0.0256***	(-21.08)	-0.0056***	(-4.15)				
Balance of Payments (% GDP)	-42.082***	(-16.28)	-7.644***	(-28.39)				
FX rate volatility					122.6***	(12.11)	36.8**	(2.97)
10-year government bond yield					25.43***	(14.63)	2.31	(0.92)
Constant	Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes	
Firm-year observations	1,741,707		1,741,707		178,618		178,618	
Firms	429,978		429,978		48,470		48,470	
Distressed firms	30,900		30,900		4,278		4,278	
Log likelihood	-127,673.50		-59,226.49		-15,147.90		-5,004.10	
Wald test	55,783.8***		57,385.55***		8,083.9***		9,835.47***	

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

Table 2.9

Sub-periods' Analysis (8 countries)

Panel A. Estimation results

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

	2002-2006		2003-2007		2004-2008		2005-2009		2000-2009	
Earnings before taxes to total assets	-0.819***	(-9.21)	-0.764***	(-9.02)	-0.824***	(-11.07)	-0.757***	(-14.16)	-0.763***	(-15.53)
EBITDA to interest expenses	-0.0000248***	(-5.91)	-0.0000390***	(-9.78)	-0.0000477***	(-12.25)	-0.0000544***	(-15.24)	-0.0000451***	(-14.58)
Current liabilities to total assets	1.789***	(68.72)	1.684***	(74.63)	1.530***	(75.25)	1.379***	(92.63)	1.417***	(102.97)
Cash flow to current liabilities	-0.557***	(-5.66)	-0.635***	(-6.49)	-0.493***	(-5.87)	-0.523***	(-9.10)	-0.491***	(-9.16)
Turnover to total liabilities	-0.0900***	(-13.08)	-0.0983***	(-14.83)	-0.118***	(-18.17)	-0.169***	(-31.22)	-0.176***	(-35.56)
Size (ln(totals assets))	-0.0980***	(-13.25)	-0.0316***	(-4.85)	0.0446***	(7.34)	-0.0188***	(-4.01)	-0.0913***	(-22.14)
Dummy equal to 1 if shareholders are more than 2	-0.245***	(-11.31)	-0.279***	(-15.03)	-0.246***	(-14.68)	-0.277***	(-20.73)	-0.272***	(-21.76)
Dummy equal to 1 if SME is located in an urban area	0.125***	(4.93)	0.0757***	(3.53)	0.0959***	(5.24)	0.141***	(10.26)	0.144***	(11.01)
FX rate (% change)	-1,421.8***	(-29.32)	-1,452.5***	(-33.02)	-478.9***	(-13.02)	-541.8***	(-14.16)	-1,689.9***	(-59.01)
Unemployment	2.117***	(4.38)	-0.462	(-0.97)	2.082***	(6.89)	4.423***	(28.04)	1.914***	(12.58)
Economic sentiment indicator	-0.0169***	(-7.70)	-0.0368***	(-20.08)	-0.0106***	(-10.39)	-0.00570***	(-6.75)	-0.0258***	(-34.90)
Loans granted to non-financial sector (% change)	-6.238***	(-50.60)	-5.288***	(-48.89)	-2.347***	(-20.75)	-4.202***	(-49.85)	-4.407***	(-58.07)
Years to resolve insolvency proceedings	-0.0497***	(-5.03)	0.0520***	(8.94)	0.0981***	(23.87)	0.157***	(46.05)	0.0958***	(27.57)
Industry 1 (Agriculture, Mining, Manufacturing)	0.0628*	(2.06)	0.0211	(0.80)	0.0915***	(3.75)	0.0712***	(3.82)		
Industry 2 (Utilities, Transportation, Communication)	-0.186***	(-4.13)	-0.0976**	(-2.65)	0.0245	(0.75)	0.0112	(0.46)	-0.0762***	(-3.56)
Industry 3 (Construction)	-0.193***	(-6.07)	-0.0267	(-0.99)	0.179***	(7.32)	0.218***	(11.78)	0.0798***	(5.84)
Industry 4 (Trade)	-0.0571*	(-1.99)	-0.113***	(-4.56)	-0.0202	(-0.88)	0.0265	(1.56)	-0.0295*	(-2.50)
Industry 5 (Accommodation and Food)	0.264***	(5.57)	0.156***	(4.06)	0.190***	(5.82)	0.226***	(9.57)	0.212***	(10.18)
Constant	Yes		Yes		Yes		Yes		Yes	
Duration	Yes		Yes		Yes		Yes		Yes	
Firm-year observations	1,079,429		1,367,406		1,704,810		2,056,890		2,721,861	
Firms	385,546		637,299		646,812		636,008		644,234	
Distressed firms	15,914		20,665		24,276		42,351		49,355	
Pseudo R-squared	0.200		0.158		0.150		0.125		0.171	
Log likelihood	-58,989.5		-69,826.7		-91,025.0		-111,420.9		-204,538.30	
Wald test	23,784.5***		31,818.7***		39,646.7***		68,451.6***		84,526.8***	

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

Table 2.9. Cont.

	2002-2006	2003-2007	2004-2008	2005-2009	2000-2009
	Panel B. Performance over next year				
Hosmer-Lemeshow decile					
1 to 5	8.05%	11.84%	11.94%	-	-
8	14.43%	17.56%	13.18%	-	-
9	19.67%	20.34%	18.99%	-	-
10	44.06%	36.30%	40.75%	-	-
8 to 10	**78.15%**	**74.20%**	**72.93%**	**-**	**-**
Area under the ROC curve	0.818	0.783	0.796	-	-
	Panel C. Performance over last year (2009)				
Hosmer-Lemeshow decile					
1 to 5	12.80%	12.25%	11.94%	-	-
8	16.34%	16.19%	13.18%	-	-
9	18.45%	18.91%	18.99%	-	-
10	35.86%	36.82%	40.75%	-	-
8 to 10	**70.65%**	**71.93%**	**72.93%**	**-**	**-**
Area under the ROC curve	0.780	0.783	0.796	-	-

2.8 Figures of Chapter 2

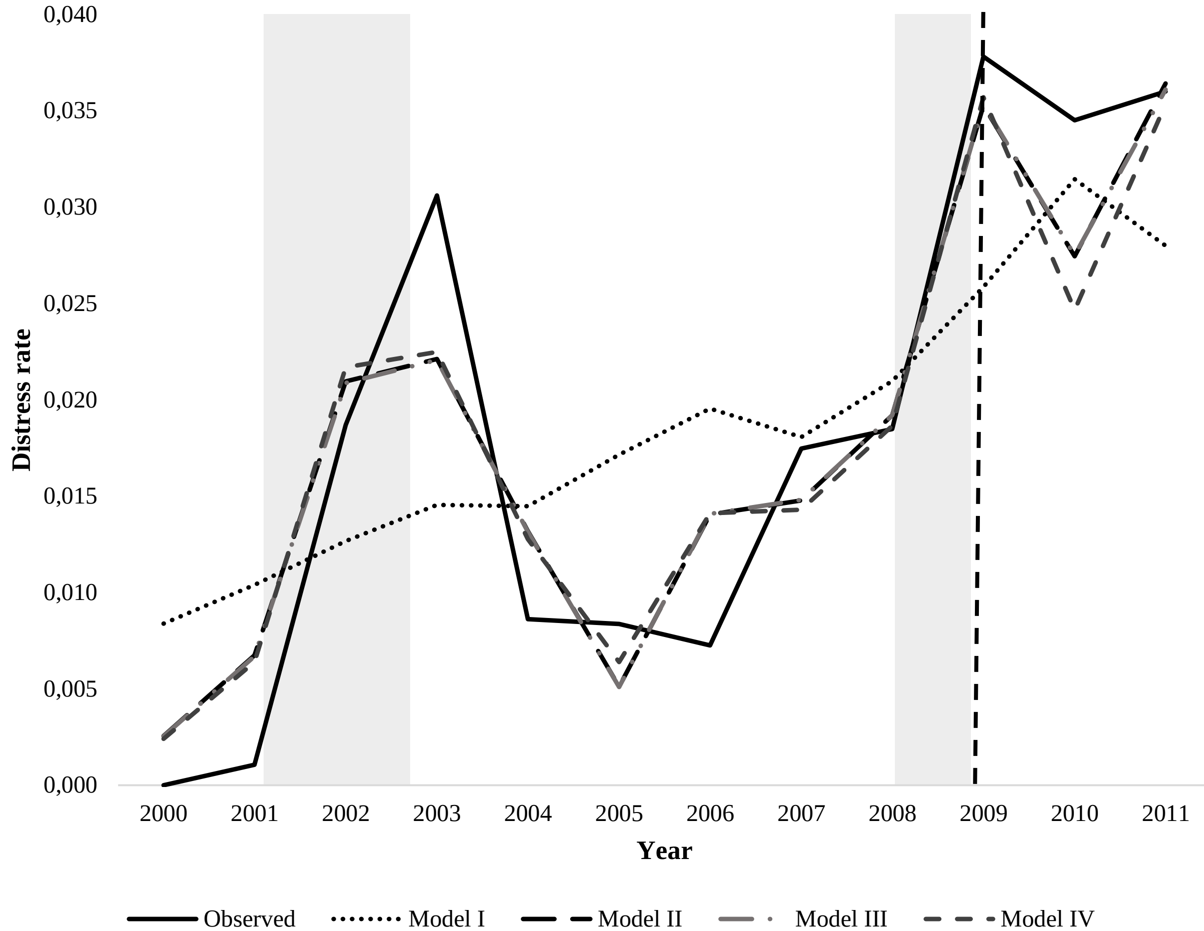

Figure 2.1. Predicted and Observed Distress Rate. The figure plots the predicted distress rate based on models I to IV of Table 2.5, along with the observed distress rate, for the period 2000 to 2011. Since we have financial but not distress information for year 2010 and for only a few firms for 2011, we do the following: (i) we use the estimated coefficients from 2000-2009 to predict the distress rate for 2000-2011; (ii) we use the in-sample observed distress rate for 2000-2009 and we obtain the observed distress rate for 2010-2011 from Creditreform. The columns denote recession periods in the euro area, as indicated by OECD. The vertical dashed line separates in-sample (2000-2009) and out-of-sample (2010-2011) periods.

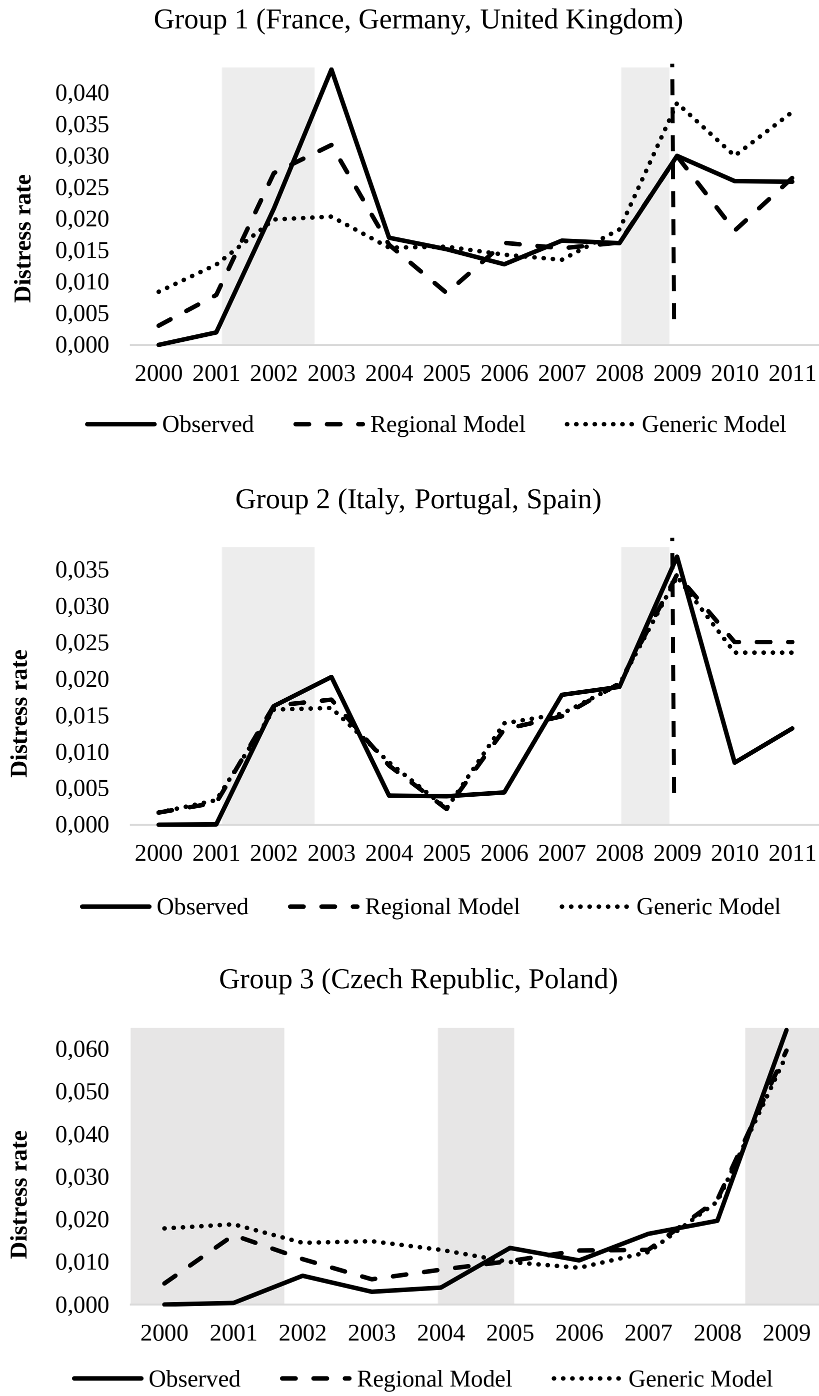

Figure 2.2. Predicted and Observed Distress Rates for the 3 groups. The figure plots the predicted distress rates based on the regional and generic models of Table 2.6, along with the observed distress rate for each group. The predicted distress rate is the simple average of the probabilities of distress of all firms in each group and period. Since we have financial but not

distress information for 2010 and for only a few firms for 2011, we do the following: (i) use the estimated coefficients from 2000-2009 to predict the distress rate for 2000-2011; (ii) use the in-sample observed distress rate for 2000-2009 and obtain the observed distress rate for 2010-2011 from Creditreform. Creditreform does not provide distress information for group 3. For groups 1 and 2, the columns denote recession periods in the euro area, and for group 3, recession periods in the Czech Republic and Poland, as indicated by OECD. When present, the vertical dashed line separates in-sample (2000-2009) and out-of-sample (2010-2011) periods.

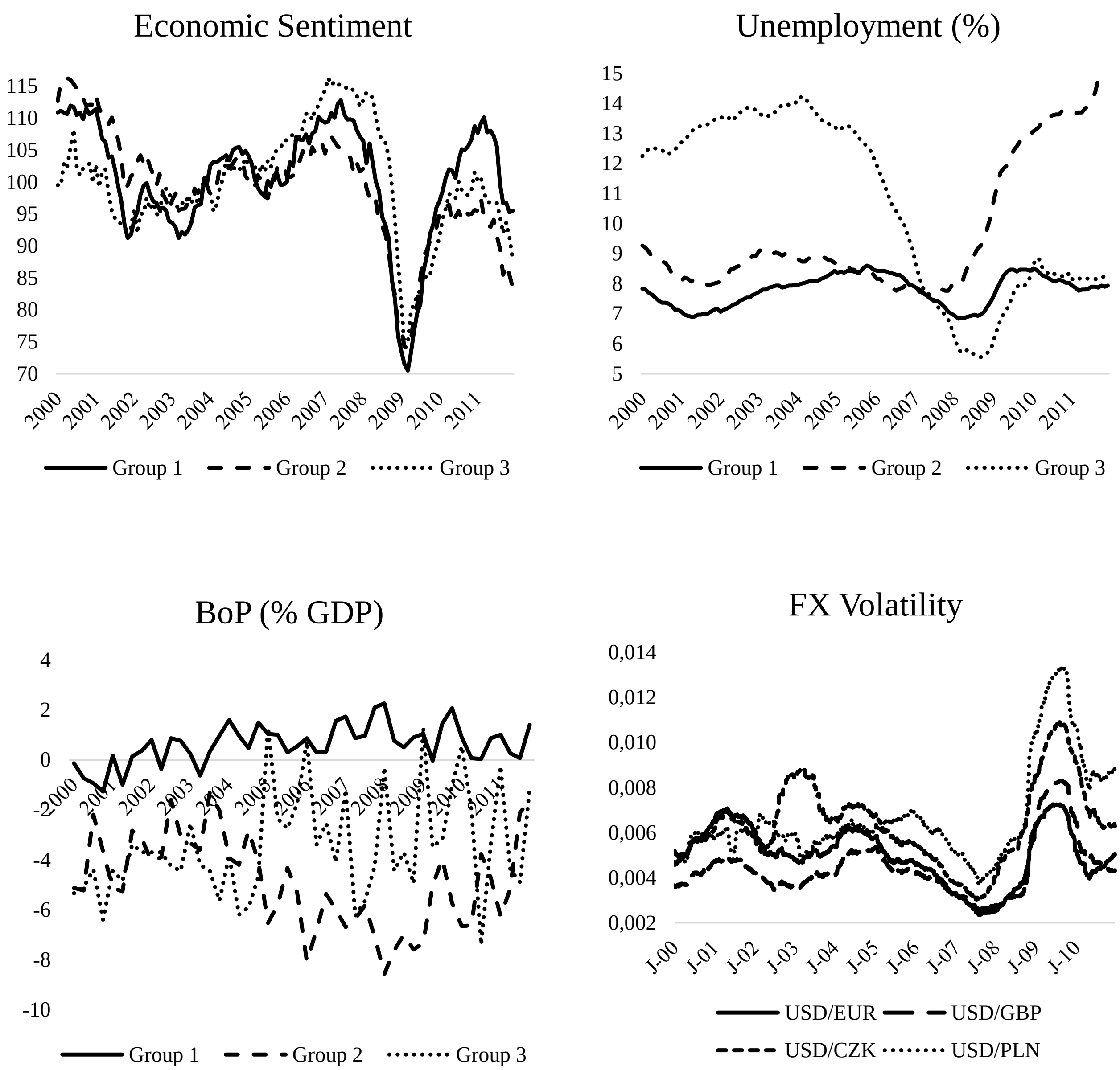

Figure 2.3. Macroeconomic variables 2000-2011. The figure plots the aggregate time series for four macroeconomic variables. The economic sentiment indicator and unemployment values are rolling annual averages at monthly frequency. Balance of payments values are rolling annual averages at quarterly frequency. Foreign exchange rate volatility values are rolling annual averages at daily frequency.

2.9 Appendices of Chapter 2

Appendix 2.1. List of Systematic Variables

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

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

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

Appendix 2.2. Insolvency Regimes

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

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

Appendix 2.3
Distress Statistics using Main and Alternative Distress Definitions

The appendix summarizes the properties of our main and alternative distress indicators for the overall sample, the three regional sub-samples and the eight countries. It gives the number of total SMEs and distressed SMEs and the distress rate. According to our main distress definition, a firm-year is distressed if the following two conditions are both met: (i) it is the last firm-year for which we have available financial statements before the firm leaves the sample; (ii) the firm (a) is either in default, in receivership, bankrupt, or in liquidation or (b) it has no updated status information and disappears from the sample before 2010 with negative equity in the last year. In our alternative distress definition, we exclude all firms that disappear from the sample before 2010 without updated status situation. Under our alternative distress definition, the sample decreases by 41% and includes only distress incidents that are strictly linked with a legal insolvency procedure.

	Firm-years		Firms		Distressed		(%) firm-years		(%) firms	
	Main	Alternative	Main	Alternative	Main	Alternative	Main	Alternative	Main	Alternative
Panel A. Overall sample and regional subsamples										
Overall sample	2,721,861	1,594,433	644,234	389,347	49,355	12,362	1.81	0.78	7.66	3.18
Group 1	801,536	332,547	165,786	66,306	14,177	5,646	1.77	1.70	8.55	8.52
Group 2	1,741,707	1,185,258	429,978	302,959	30,900	6,338	1.77	0.53	7.19	2.09
Group 3	178,618	76,628	48,470	20,082	4,278	378	2.40	0.49	8.83	1.88
Panel B. Countries										
Germany	21,681	5,322	5,954	1,326	319	8	1.47	0.15	5.36	0.60
France	724,060	309,230	145,918	61,030	12,222	5,353	1.69	1.73	8.38	8.77
United Kingdom	55,795	17,995	13,914	3,950	1,636	285	2.93	1.58	11.76	7.22
Italy	278,630	209,924	89,666	71,348	2,257	219	0.81	0.10	2.52	0.31
Portugal	487,664	402,898	148,645	123,193	10,396	3,702	2.13	0.92	6.99	3.01
Spain	975,413	572,436	191,667	108,418	18,247	2,417	1.87	0.42	9.52	2.23
Czech Republic	119,677	59,856	33,305	16,804	3,014	244	2.52	0.41	9.05	1.45
Poland	58,941	16,772	15,165	3,278	1,264	134	2.14	0.80	8.33	4.09

Chapter 3 Pricing Default Risk: The Good, The Bad, and The Anomaly

3.1 Introduction

Finance theory suggests that if default risk is systematic (and thus non-diversifiable) it should be positively correlated with stock returns in the cross-section of firms. However, in the empirical literature there are two main strands that deliver contradictory findings regarding the sign and significance of this relationship. On the one hand, Vassalou and Xing (2004) and Chava and Purnanandam (2010) document a positive relationship between default risk and stock returns in the US and Aretz, Florackis and Kostakis (2014), in a recent working paper, report similar findings using an international sample. On the other hand, several studies find a negative relationship between default risk and returns, the so-called "default anomaly". Examples are Dichev (1998), Griffin and Lemmon (2002), Campbell, Hilscher and Szilagyi (2008), Garlappi, Shu and Yan (2008), Avramov et al. (2009), Da and Gao (2010), Garlappi and Yan (2011), and Conrad, Kapadia, and Xing (2012) in the US, Bauer and Agarwal (2014) in the UK and Gao, Parsons and Shen (2013) internationally.[8]

[8] Some of the explanations offered for this puzzling evidence are: (i) violations of the absolute priority rule (Garlappi, Shu and Yan, 2008; Garlappi and Yan, 2011): higher shareholder bargaining power reduces the risk of the shareholders' residual claim, thus returns close to default; (ii) long-run risk (Avramov, Cederburg, and Hore, 2011): firms close to default are less exposed to long-run risk because they are not expected to survive for long, and hence have lower returns; (iii) glory (Conrad, Kapadia, and Xing, 2012): firms with high default risk are glory stocks that realize high returns in the future, so their current low returns are not a good estimate of their future returns. (iv) psychological reasons (Gao, Parsons and Shen, 2013): investors are overconfident about high default risk stocks, keeping their prices high and subsequently leading to sudden corrections and low returns; (v) neglected

These literature strands focus on the firm's physical probability of default (PD) as a measure of default risk. In most cases, they use either market-based PDs (calculated under Merton's (1974) framework) or accounting-based PDs (such as Altman's Z-score, Ohlson's O-score, and the popular measure used by Campbell, Hilscher and Szilagyi (2008)). Hence, these studies implicitly assume that physical PDs are monotonically related to risk-neutral PDs and that, as physical PDs increase, so does the exposure to aggregate default risk. However, George and Hwang (2010) argue that a firm's physical PD does not necessarily reflect its systematic risk. In a theoretical model, they show that firms with high SDR exposure choose low leverage levels, which in turn lowers their physical PDs, therefore creating a negative relationship between PDs and returns. In the same spirit, Kapadia (2011) finds that firms with high physical PDs do not co-vary with aggregate distress, suggesting that the low returns of high PD stocks are not due to exposure to aggregate distress. Similarly, Avramov, Cederburg and Hore (2011) show that firms with high idiosyncratic volatility (often identified as firms with high PDs) have low SDR exposure and low returns, thus suggesting a link between idiosyncratic volatility and default anomalies.[9]

Following George and Hwang's (2010) and Kapadia's (2011) influential work, many recent working papers use proxies of risk-neutral PDs instead of physical PDs to measure default risk, and most document a positive relationship between default risk and returns. Examples are Chan-Lau (2006), Nielsen (2013) and Friedwald, Wagner and Zechner (2014), who use credit default swap (CDS) spreads, and Anginer and Yildizhan (2014), who calculate credit risk premia from

profitability (Bauer and Agarwal, 2014): distress risk without profitability related information is not relevant in pricing.

[9] Other studies that document a negative relationship between idiosyncratic volatility and stock returns (the IV anomaly) include Ang et al. (2006) and Barinov (2012).

corporate bond spreads to proxy for risk-neutral PDs. The disadvantage of these studies is that they can only calculate risk-neutral PDs for firms that have CDS or bond information available. These firms constitute a small fraction of total firms and are usually the largest ones. For example, Ozdagli (2013) argues that CDS data are available for only about 20% of US public firms (and are reliable only after 2004).

Of the above studies that focus either on CDS or bond data, only Anginer and Yildizhan (2014) extend their analysis to a larger sample of firms for robustness purposes. To do this, they use physical PDs of US firms with CRSP-COMPUSTAT data available and calculate sensitivities of these PDs to the median PD in their sample, which they use as a proxy for aggregate default risk. Interestingly, they document a positive relationship between these sensitivities and stock returns. Our study is close to their analysis. Specifically, we build on this methodology, which was introduced by Hilscher and Wilson (2013), and extend Anginer and Yildizhan (2014) in three ways that we describe below.

First, we use as a measure for aggregate default risk the CBOE Volatility Index (VIX) instead of the median PD. VIX is a good proxy for aggregate default risk since it is positively correlated with credit spreads, as documented in the literature on CDS (Pan and Singleton, 2008 (distress risk premium)) and corporate bonds (Collin-Dufresne, Goldstein, and Martin, 2001; Schaefer and Strebulaev, 2008). Moreover, VIX is strongly correlated with European volatility indices (correlations higher than 0.90), which are generally available only from 2000 onwards. Several studies also connect VIX with stock returns. Ang et al. (2006) calculate the sensitivity of individual returns to changes in VIX, and show that firms that perform well when VIX increases experience low average returns because they are a hedge against market downside risk. Barinov (2012) additionally shows that both firms with very negative and very positive return sensitivities to VIX changes are smaller and have higher BM ratios. In unreported results, we also use the median PD as an alternative proxy for aggregate default risk and all our results remain robust. However, in our large sample of very heterogeneous countries, the median PD is a rather noisy measure and demonstrates higher auto-correlation than monthly VIX. Thus, we believe VIX is a more appropriate measure and further motivates its use in Section 3.4.

Second, instead of focusing on the US market, which has already been largely explored, we study a comprehensive sample of European firms from 22 countries, which notably also includes smaller firms. These firms are often neglected, but constitute the vast majority of firms listed on European exchanges. This heterogeneity is important as previous work has often associated default risk to other firm characteristics (such as size and book-to-market ratios). Thus, the inclusion of small stocks allows us to reconcile our findings with these earlier results.

Finally, we break down the physical PDs into systematic and idiosyncratic components and study the relationship between returns and the two components of physical PD separately. This enables us to detect the origin of the default anomaly. We refer to the systematic component as systematic default risk (SDR) beta and to the idiosyncratic component as idiosyncratic default risk (IDR) alpha. Specifically, we sort the stocks in our sample on both SDR betas and IDR alphas instead of only SDR betas (as Anginer and Yildizhan do) and perform several double-sorts in order to better identify the source of the anomaly and enforce our statements.

Our study is also close to Gao, Parson and Shen (2013) and Aretz, Florackis and Kostakis (2014), who study the relationship between default risk and stock returns in international samples. Gao, Parson and Shen (2013) study 39 countries and document a negative relationship between stock returns and default risk, as measured by Moody's Expected Default Frequency (EDF) measure. They provide a behavioral explanation for the anomaly that sees investors being overconfident about high default risk stocks, keeping their prices high and subsequently leading to sudden corrections and low returns. Contrary to their findings, Aretz, Florackis and Kostakis (2014) study 14 developed markets using the accounting-based measure of Campbell, Hilsher and Szilagy (2008) and document a positive relationship between stock returns and default risk.

Thus, Gao, Parson and Shen (2013) use market-based PDs and Aretz, Florackis and Kostakis (2014) use accounting-based PDs to proxy for SDR exposure. But, as we already discuss above, these physical PDs are not necessarily good measures of such exposure. Our study differs from the these two as we use the simple and intuitive method described above to break these physical PDs down into systematic and idiosyncratic components. We then study the relationship between returns and these two components separately in order to better capture how exposure to aggregate default risk is priced.

Our main hypothesis, which we confirm empirically, is that stocks with high sensitivities of their PDs to VIX (not necessarily high PDs per se) will have higher average returns, because investors will require a premium for holding such stocks. Therefore, the documented default anomaly in the literature is only the result of incorrect measurement of the exposure to aggregate default risk.

The remainder of the study is organized as follows. Section 3.2 describes the data. Section 3.3 studies the relationship between the physical PDs and stock returns. Section 3.4 first describes the method to break down the physical PDs into systematic and idiosyncratic components, and then discusses the relationship between these different components and stock returns. Section 3.5 performs further tests to support the explanation of the default anomaly. Finally, Section 3.6 concludes.

3.2 The Data

Our study covers publicly listed firms from 22 European countries, during the period January 1990 to December 2012. We use Thomson Reuters' Datastream for market data and Thomson Reuters' Worldscope database for the firms' accounting information.

To guarantee a certain level of market exchange activity, we include in our analysis only the 22 European countries that had established exchanges on or before 1980 (for a total of 34 exchanges). We exclude years 1980-1989 due to the limited number of companies with available data. We also follow previous studies in the field and exclude financial firms (ICB[10] 8000 Financials) and firms with negative BM ratios. To reduce the influence of outliers and account for measurement errors, we exclude firms with a market capitalization below the 1st percentile for all observations. This essentially leaves in our sample firms with a market capitalization above roughly one million euros.[11] Moreover we only retain firms that have at least two years of data available, that is sufficient historical data for the calculation of physical PDs. To avoid duplicate observations, we do the following: for firms that are traded on more than one European exchange, we keep data from the market where the firm has been traded for the longest period, this is almost always the home market, and; if a firm has issued more than one type of common share, we use data of the share type that constitutes the majority of common equity.

An important feature of our database is the compiled information on default events. As the reason for delisting is not usually available in Datastream, we manually track the status of the delisted firms from other sources (such as Amadeus and Orbis Europe databases), as well as

[10] The Industry Classification Benchmark (ICB) is an industry classification taxonomy launched by Dow Jones and FTSE in 2005.

[11] US studies usually exclude stocks with prices below 5 USD, but such a condition is inappropriate for our European sample (i.e. in Europe, the median stock price is approximately 5 EUR).

various public internet sources. Therefore, we are able to identify if the delisting of a security is due to default (bankruptcy or liquidation) or other reasons (i.e. mergers). To illustrate this point, Table 3.1 reports the average number of active firms per year, as well as the number of firms that were delisted due to default each year.

Nonetheless, information on delisting is also not available in Datastream. Thus we follow Campbell, Hilscher and Szilagyi (2008) and use the last available full-month return, assuming that our portfolios sell stocks that are delisted at the end of the month before delisting. This approach gives a conservative estimate of the default anomaly. Results are qualitatively the same if we follow Vassalou and Xing (2004) and set delisting returns for stocks that default equal to -100 percent (assuming a zero recovery rate).

After applying the filters described above and merging different data sources, we are able to calculate physical PDs and draw results for a final sample of 806,157 firm-months (corresponding to 8,439 firms) across the 22 European countries. Table 3.2 characterizes this final sample with respect to the distribution of firms across size classes and countries. The average size in our sample is lower than previous US studies, because small stocks constitute the majority of traded firms in European exchanges. In terms of international breakdown, the representativeness of the different countries in our sample seems to be in line with the literature (e.g. Gao, Parsons, and Shen, 2013). Unsurprisingly, more developed markets contribute a greater share of observations to the sample, with the U.K. (32.54%), France (13.34%) and Germany (13.08%) collectively comprising more than half.

We also resort to various other public data sources. Regarding aggregate default risk proxies, we use the CBOE VIX, as well as the European volatility indices VSTOXX, VFTSE and VDAX (for EUROSTOXX 50, FTSE 100 and DAX respectively) and the European credit default swap index iTraxx. We focus on VIX in the main analysis, as it is the only index available from January 1990 onwards. The Fama-French factors SMB and HML and the market factor EMKT for Europe are obtained from Kenneth French's web page. For the risk-free rate, we use monthly observations of the 1-year T-bill, available from the Federal Reserve Board Statistics.[12]

3.3 The Physical Probabilities of Default and Stock Returns

3.3.1 Calculating Physical PDs

We follow Vassalou and Xing (2004) in calculating our main physical PD measure. As their methodology is based on the Merton model, we also refer to the estimated physical PD as the Merton measure. In order to calculate monthly PDs under this approach, we use data on current and long-term debt, as well as market capitalization for all the firms in our sample. We perform all calculations for the individual monthly PDs in local currency to minimize the effect of exchange rate volatility. Appendix 3.1 presents more details on the Merton measure, its calculation and overall performance.

Table 3.3 shows descriptive statistics for the estimated Merton measure by country. Since other firm characteristics (such as size and BM ratios) have been associated with default risk in the literature, Table 3.3 also includes descriptive statistics for these variables (along with raw average returns). Overall the results show that there is significant heterogeneity among European countries in terms of PDs, size, and BM. Markets such as Romania (16.69%) and Bulgaria (14.29%) have the highest average PDs, while other countries such as Switzerland (3.13%) and the Netherlands (3.42%) have very low average PDs.

[12] We use a US risk-free rate since we do not have a sufficiently long time series of data for the German equivalent. Similarly, Kenneth French calculates the European factors using a US risk-free rate.

Although the performance results in Appendix 3.1 suggest that the Merton measure is indeed a good default predictor, we also calculate an alternative default measure for robustness purposes. In particular, we follow Campbell, Hilscher and Szilagyi (2008) in calculating a physical PD measure using a multi-period logit regression framework. We refer to this alternative PD as the CHS measure. We are able to calculate the CHS measure for 755,243 firm-months (7,980 firms). For more details on the methodology, please refer to Appendix 3.2.

Figure 3.1 summarizes the results. In Panel A, we plot the monthly aggregate Merton and CHS measures for firms in the overall sample (defined as simple averages of the values of all firms). The two PD measures have a very high correlation of 0.92, but their magnitude is different since the CHS measure produces lower PDs than the Merton measure. The columns in the plot denote recession periods in the euro area (as indicated by the OECD), so we can also observe that both measures vary greatly with the business cycle and increase during downturns. Panel B plots the monthly aggregate Merton measure and values of the volatility index VIX at the end of each month. It is again apparent that Merton PDs and VIX co-move closely together throughout the economic cycle. Both are higher during recessions, when economic theory suggests that the stochastic discount factor is high. This finding provides initial evidence that VIX captures aggregate default risk information.

For reasons of brevity and given the high time-series and cross-sectional correlations between the two PD measures, we present results only with the estimated Merton measure. We justify this choice in two ways. Firstly, the CHS measure may suffer from an in-sample bias, since we use data from the whole sample period to estimate PDs. Secondly, we are able to estimate the CHS measure for a smaller sample of firms compared to the Merton case. Nonetheless, all our results are robust to the choice of physical PD measure.

3.3.2 The Default Anomaly: Physical PDs and Stock Returns

As a first part of our analysis, we study the possible existence of a default anomaly in Europe. In particular, we explore the cross-sectional relationship between stock returns and default risk by conducting portfolio sorts on the physical PDs.

Each month, from January 1990 to December 2012, we use the most recent PD for each firm and sort the stocks into five portfolios. [13] To account for possible country effects (concentration of risky stocks in certain countries and/or accounting differences), we follow an approach similar to Lewellen (1999) and Barry et al. (2002): at the beginning of each month, we adjust the available PDs from stocks in the overall sample by the average country PD for this month. Then we sort all stocks into portfolios based on the adjusted PDs. We perform similar adjustments for all the tests that follow but our results remain robust if we do not adjust for country averages. If the integration amongst European markets is high, it is not necessary to perform such adjustments. Nevertheless, our sample consists of 22 European countries, of which three are not members of the European Union, thus it is not very plausible to assume a very high degree of integration.

Table 3.4 displays the results. In Panel A, we report both equally and value-weighted monthly raw returns and alphas (excess risk-adjusted returns) of the five portfolios. We also construct high-low portfolios (which are long the highest PD stock quintile and short the lowest PD stock quintile) and report raw returns and alphas for these portfolios (the alphas are obtained using the factor-mimicking portfolios for Europe available on Kenneth French’s website). The results show that the difference in returns between high and low PD stocks is almost always

[13] All results are qualitatively similar using ten portfolios instead of five.

negative, in line with the literature that documents a possible default anomaly. This relationship is almost monotonic, but differences are not always significant. Thus, there is weak evidence that the high PD stocks earn on average lower returns than the low PD stocks, though this under-performance does not demonstrate significance.

In Panel B of Table 3.4, we report the estimated factor loadings for excess, equally and value-weighted returns on the four Fama-French-Carhart factors. We find that high PD portfolios have higher loadings on the market factor (EMKT), the size factor (SMB) and the value factor (HML). This shows the prevalence of small and value stocks in the high PD portfolios. To complement this analysis, in Panel C we report some relevant characteristics of the five portfolios. Average size decreases monotonically across the portfolios, and average BM increases monotonically, again reflecting the dominance of small and high BM firms among the high PD stocks. The high PD stocks also have high leverage ratios (LRs) and, in accordance with Chen and Zhang (2010), low return on assets (ROAs) ratios.

3.4 Understanding Default Effects

3.4.1 Decomposing the Physical PDs into Systematic and Idiosyncratic Components

Our findings in the previous section appear to support the existence of a default anomaly, since an investing strategy that buys the highest PD stocks and shorts the lowest PD stocks has, on average, negative returns. At first glance, these results suggest that default risk is, at best, not priced into the cross-section of stock returns. However finance theory suggests that, only if default risk is systematic and thus non-diversifiable, it should be positively correlated with expected stock returns. In other words, investors demand a premium to hold the stocks of firms with high exposure to aggregate default risk, not necessarily firms with high physical PDs. In fact, George and Hwang (2010) argue that a firm's physical PD does not necessarily reflect its SDR exposure. Therefore in this part, we break down the physical PDs into systematic and idiosyncratic components, and investigate empirically if the physical PDs are a good measure of firm exposure to aggregate default risk.

3.4.2 The Methodology

To calculate SDR exposure, we follow an approach similar to Hilscher and Wilson (2013) and Anginer and Yildizhan (2014), by assuming that a firm's PD is exposed to a single common factor. This factor is the aggregate default risk. Therefore the firm's SDR exposure is measured as the sensitivity of its PD to this factor (we refer to this sensitivity as the SDR beta). To compute monthly SDR betas for all firms in our sample we estimate the following regression for each firm over 24-months rolling windows:

$$PD_{i,t} = \alpha_i^{IDR} + \beta_i^{SDR} X_t + \varepsilon_{i,t} \ , \qquad (1)$$

where $PD_{i,t}$ is the physical PD for firm i in month t, X_t is the aggregate default risk measure, α_i^{IDR} is the IDR alpha and β_i^{SDR} is the SDR beta for firm i in month t, obtained from the rolling regressions method.[14] To avoid auto-correlation concerns we estimate our regressions using both changes and levels, and results remain robust. We are able to calculate SDR betas and IDR alphas for 624,084 firm-months (7,140 firms) for the period from January 1992 to December

[14] The specification in (1) does not of itself constrain the PD to lie between zero and one. Hilscher and Wilson (2013) argue that this is not a problem, as long as most of the estimated PDs are small (so that $P(1-P) \approx P$). Our estimated PDs satisfy this condition. Also, since the percentage of the variability explained by the residuals in our regressions is small, our results do not change if we include the error term in the idiosyncratic component of default risk.

2012. The sample is smaller than before because we need two years of PD history for the estimation. Essentially, we cannot calculate SDR betas for January 1990 to December 1991.

3.4.3 VIX and Aggregate Default Risk

As a proxy for aggregate default risk, we use the volatility index VIX. Several authors study the relationship of equity returns and VIX (Chira et al., 2013). We are also not the first to link default risk with VIX. Numerous studies find VIX to be an important indicator of credit spreads, as shown in the literature on CDS (Pan and Singleton, 2008) and on corporate bonds (Collin-Dufresne, Goldstein, and Martin, 2001; Schaefer and Strebulaev, 2008). Table 3.5 motivates further the use of VIX in our empirical analysis. Panel A presents summary statistics for VIX and its monthly change, Δ_mVIX. Panel B reports the highly positive correlation coefficients between VIX and three European volatility indices, which suggests that VIX successfully captures aggregate volatility in Europe. Panel C reports the negative correlation coefficients between Δ_mVIX and the monthly change of two widely used European stock indices, EUROSTOXX 50 and MSCI Europe. This finding is in line with the theoretical model of Bansal et al. (2014), according to which stock returns have, on average, negative volatility betas. Panel D reports the negative correlation coefficients of Δ_mVIX with EMKT and SMB, which is in accordance with Ang et al. (2006). For HML, the correlation is very low. Lastly, the regression results of Panel E show that VIX can explain a substantial portion of time-variation in both the aggregate and the median physical PD. In unreported results, we follow the US studies of Hilscher and Wilson (2013) and Anginer and Yıldızhan (2014) and use the median PD as an alternative proxy for aggregate default risk. Hilscher and Wilson (2013) find that the median PD is highly correlated with the first principal component which explains the majority of variation in

PDs across ratings. All results remain robust, however, in our large sample of very heterogeneous countries, the median PD is a rather noisy measure. It also demonstrates higher auto-correlation than monthly VIX.

3.4.4 Physical PDs, Systematic Betas, and Idiosyncratic Alphas

In Panel C of Table 3.4, we show that stocks in the highest PD quintile have relatively low SDR betas, whereas their IDR alphas are very high. This empirical finding is in accordance with the theoretical model of George and Hwang (2010) and suggests that the physical PDs may not be a good measure of firm exposure to aggregate default risk. We now turn to the analysis of the relationships between stocks returns and the two components of the PD separately.

3.4.4.1 SDR Betas and Stock Returns: A Premium on Exposure to Aggregate Default Risk

To examine if exposure to aggregate default risk are rewarded in the cross-section of stock returns, we repeat the portfolio analysis of Section 3.3. using the SDR betas as the sorting variable. Each month, from January 1992 to December 2012, we use the most recent SDR beta for each firm and sort the stocks into five portfolios. As before, we adjust monthly SDR betas by their monthly country average. Table 3.6 reports the results.

Panel A shows that the difference in returns between high and low SDR beta stocks is now always positive for both equally and value-weighted returns and significant in the case of equally-weighted returns. A portfolio strategy buying the highest SDR beta quintile and shorting the lowest SDR beta quintile of stocks gives an equally-weighted four-factor alpha of 0.33 percent monthly (4.01 percent annually), significant at a five percent level. The positive relationship between returns and SDR betas is almost always monotonic. Thus, when we use an

SDR measure to sort the stocks, there is evidence of a positive relationship between default risk and returns, in line with theoretical models.

In Panel B, we see that factor loadings on the market factor (EMKT) and the size factor (SMB) do not decrease monotonically along the SDR beta portfolios. Specifically, both high and low SDR beta stocks have higher loadings than medium SDR beta stocks. This indicates that small stocks are not homogeneous with respect to their SDR exposure. The factor loadings on the value factor (HML) are mostly insignificant. These results suggest that our SDR measure conveys information that is not captured by traditional risk factors.

Panel C reports some characteristics of the portfolios. First, SDR betas exhibit large cross-sectional dispersion, ranging from -0.62 to 0.89, indicating that the effect of aggregate default risk varies substantially across stocks. In accordance with Barinov (2012), negative SDR betas indicate that these portfolios are indeed a hedge against increases in VIX, which justifies their low returns. Second, we find interesting non-monotonic patterns across the beta portfolios: (i) both high and low SDR beta stocks have higher PDs than medium SDR beta stocks; (ii) they also have higher LRs and lower ROAs; (iii) they are also, on average, smaller in size and have higher BM ratios (which is consistent with the results from portfolio sorts on credit risk premia estimated from CDS spreads by Friedwald, Wagner and Zechner, 2014). Therefore the SDR beta conveys information that is different from that incorporated in other common default risk measures and stock characteristics. Finally, we find a negative relationship between SDR betas and IDR alphas, as the idiosyncratic component of the PD increases almost monotonically across the SDR beta portfolios. This is in accordance with Avramov et al. (2013), who document a negative cross-sectional relationship between exposure to systematic and firm-specific risks.

In unreported results, we also test for return persistence in our SDR beta sorted portfolios. Da and Gao (2010) argue that the high returns of risky stocks do not compensate for SDR, but the result of short-term return reversal caused by price pressure in the month of portfolio formation. Thus, in accordance with the default anomaly literature, they find that risky stocks deliver low returns if the second month after portfolio formation is used instead. We find no evidence of return reversal: the return of the highest and lowest SDR beta quintiles differ 8 months before portfolio formation, the difference is maximized in the portfolio formation month, and persists for almost 8 months after portfolio formation (even if we assume zero recovery of firms suffering default).

To conclude, the findings in this section show that SDR betas, measured as sensitivities of the physical PDs to a common aggregate default risk factor (here VIX) are positively related to stock returns and that high PD stocks can have quite different SDR betas among them.

3.4.4.2 IDR Alphas and Stock Returns: A negative relationship

We now sort stocks based on the IDR alphas.[15] Each month, from January 1992 to December 2012, we use the most recent IDR alpha for each firm and sort the stocks into five portfolios. As before, we adjust monthly IDR alphas by their monthly country average. Table 3.7 reports the results.

Panel A shows that the difference in returns between high and low IDR alpha stocks is negative for both equally and value-weighted returns, as in the case of PDs. It is also significant at a five percent level for value-weighted returns and CAPM alphas. In Panel B, we see that

[15] Our results are robust if we measure the idiosyncratic component of default risk as the sum of IDR alphas and residuals from regression (1).

factor loadings on the market factor (EMKT) and the size factor (SMB) do not decrease monotonically along the IDR alpha portfolios, but they follow the same patterns as for SDR beta portfolios. Specifically, both high and low IDR alpha stocks have higher loadings than medium IDR alpha stocks. As before, the factor loadings on the value factor (HML) are not significant. Panel C reports some characteristics of the portfolios. IDR alphas exhibit large cross-sectional dispersion, ranging from -8.5594 to 22.5424. In accordance with our previous findings on SDR beta portfolios, both high and low IDR alpha stocks have higher PDs, are smaller, have higher BM and LRs, and lower ROA than medium IDR alpha stocks. As before, we document a negative relationship between SDR betas and IDR alphas. Therefore, stocks that have low exposure to aggregate default risk are associated with high firm-specific risks. These results are initial evidence that the default anomaly can be explained by the non-monotonic relationship between the physical PD and its systematic and idiosyncratic components.

In Tables 3.6 and 3.7, we notice that for stocks sorted on SDR betas, the equally-weighted positive returns are significant, whereas for stocks sorted on IDR alphas, the value-weighted negative returns are significant. In the case of SDR beta sorted portfolios (Table 3.6), the smaller stocks in the highest SDR beta quintile have, on average, higher returns and SDR betas, compared to the relatively larger stocks in the same portfolio. In the case of IDR alpha sorted portfolios, the smaller stocks in the highest IDR alpha portfolio have, on average, higher returns and lower IDR alphas, compared to the relatively larger stocks in this portfolio. These findings are consistent with the notion that the size effect is a proxy for default risk. They can also provide some evidence why Campbell, Hilscher and Szilagyi (2008) and other studies that have relatively larger stocks in their samples and calculate only value-weighted returns document a significant default anomaly.

3.5 Explaining the Default Anomaly

This section sheds more light on the relationship between default risk and stock returns. Our main focus is to understand what the main drivers of the default anomaly are, and therefore we apply a sequential two-sort procedure to investigate it. We use tertiles instead of quintiles to guarantee an adequate number of stocks in each portfolio (at least twenty each month). For brevity, we report value-weighted returns but results remain qualitatively similar for equally-weighted returns.

Table 3.8 examines the default anomaly in SDR beta-sorted portfolios. Specifically, each month, we first sort stocks into three portfolios based on their country-adjusted SDR beta and, within each SDR beta portfolio, we further sort stocks in three portfolios, based on the country-adjusted PD. Panel A shows the time-series monthly average of the value-weighted returns and alphas, as well as average monthly returns and alphas for portfolios that are long the highest PD tertile and short the lowest PD tertile of stocks. As expected, the default anomaly is no longer significant when we control for exposure to aggregate default risk. Panel B reports various characteristics of each portfolio. Both stocks in the highest and lowest SDR beta tertiles have higher PDs than stocks in the medium SDR beta tertiles. Still, low SDR beta stocks have lower PD levels than high SDR beta stocks. They also differ in terms of their IDR alphas. While stocks in the highest SDR beta tertiles have, on average, negative IDR alphas, stocks in the lowest SDR beta tertiles have high IDR alphas. These stocks are a hedge against aggregate default risk (which explains their low returns). Finally, size and ROA decrease and BM and LR increase monotonically as PD increases in all three SDR beta tertiles, indicating that stocks with high PDs are, on average small, value stocks, with high leverage and low profitability.

Table 3.9 examines the default anomaly in IDR alpha-sorted portfolios. The double-sorting procedure is similar to the one we describe above. Interestingly, Panel A shows that the default anomaly is significant only for stocks in the highest IDR alpha tertile, but it is absent in the other two IDR alpha tertiles. Thus, the difference in returns between high and low PD portfolios is negative and significant only when the idiosyncratic component of the PD is very high. Panel B of Table 3.9 shows very similar patterns to Table 3.8. High IDR alpha stocks have, on average, negative exposure to aggregate default risk, thus constituting a hedge against bad market conditions. Moreover, among high IDR alpha stocks, this hedging ability increases as PD increases (i.e. the SDR betas become more negative). Another interesting finding is that, in the lowest IDR alpha tertile, as PD increases, SDR betas rise and IDR alphas fall. This shows that, for stocks with low idiosyncratic risk, the physical PD is a better proxy than exposure to aggregate default risk.

Overall, the results above show that (i) the so-called "default anomaly" loses its significance when we control for exposure to aggregate default risk (SDR beta), (ii) it is only found in firms with high idiosyncratic risk when we control for IDR alpha, and (iii) it is not an "anomaly", in the sense that the negative returns on the High-Low PD portfolios are compensated for by their ability to hedge.

Finally, Table 3.10 examines the relationship between SDR betas and returns in PD sorted portfolios. Each month, we first sort stocks into three portfolios based on their country-adjusted PD and, within each PD portfolio, we further sort stocks into three portfolios, based on the country-adjusted SDR beta. As mentioned above, there is a link between the negative PD - return relationship and the negative IDR alpha - return relationship. Therefore, controlling for the PD, helps uncovering the positive relationship between SDR betas and returns and illustrates better the difference between SDR betas and PDs. In Panel A of Table 3.10, we find that the difference in returns between high and low SDR beta stocks is positive and significant in the tertile of stocks with the lowest PD. In the other two PD tertiles, there is no particular relationship between SDR beta and returns. This may happen because stocks with very high PDs are subject to market imperfections that can influence their returns, such as large arbitrage costs (Cambell, Hilsher and Svilagyi, 2008) and the divergence of opinions (Miller, 1977). Thus, exposure to aggregate default risk is significantly rewarded for stocks with low PDs, which are stocks with low market imperfections. In Panel B, we see that stocks in the lowest PD tertile (where exposure to aggregate default risk is significantly rewarded) have lower IDR alphas, are larger, have lower BM and LRs (in accordance with George and Hwang, 2010) and higher ROAs (in accordance with Chen and Zhang, 2010) compared to stocks in the other two PD tertiles.

3.6 Concluding Remarks

In this paper, we shed more light on the recent contradictory literature that explores the relationship between default risk and stock returns. We follow a simple and intuitive approach to break down physical PDs into systematic and idiosyncratic components, use the VIX as a measure of aggregate default risk and provide European evidence to study the default anomaly.

Initially, we sort stocks into quintile portfolios based on their physical PDs and, in line with the literature that documents a default anomaly, we find that the difference in returns between high and low PD stocks is negative and that the returns almost monotonically decrease as the PD increases. However, a closer look shows that the physical PD is usually a poor measure of exposure to aggregate default risk. In accordance with George and Hwang (2010), we find that stocks in the highest PD quintile have relatively low SDR exposure. We then sort stocks into quintile portfolios based on their SDR betas instead; as expected, we find a positive and significant relationship between this measure of default risk and returns. In other words, investors indeed require a premium to hold stocks that are riskier when aggregate default risk is higher. Interestingly, there are non-monotonic patterns across the SDR beta portfolios. On average, the firms in the low and high SDR beta portfolios are smaller, have higher BM, and higher physical PDs than the firms in medium SDR beta portfolios. We find that the SDR betas are negatively related to the idiosyncratic component (measured by the IDR alphas). Therefore it is the idiosyncratic (not the systematic part) driving the default anomaly. We confirm this conjecture by showing that stocks sorted on IDR alphas have, on average, lower returns. Investors do not require compensation to hold stocks with high firm-specific risk because these stocks are a source of portfolio risk diversification. Further analysis with double-sorted portfolios helps us confirm these statements.

Our results suggest that riskier stocks, as measured by the physical PDs, will tend to under-perform because they have, on average, lower exposure to aggregate default risk. Their riskiness is mostly idiosyncratic and can be diversified away, thus providing an explanation for the default anomaly typically found in the literature. On the contrary, it is the systematic component of default risk, measured by the SDR betas, that requires a return premium.

3.7 Tables of Chapter 3

Table 3.1

Defaulted Firms as a Percentage of Total Firms

The table lists the total number of active firms and delistings due to default (bankruptcy or liquidation) for every year of our sample period. The number of active firms is the average number of firms across all months of the year. The number of firms that were delisted due to default is hand-collected data from various public sources.

Year	Active Firms	Defaults	(%)
1990	1,244	1	0.08
1991	1,681	4	0.24
1992	2,072	12	0.58
1993	2,242	6	0.27
1994	2,322	9	0.39
1995	2,374	11	0.46
1996	2,398	14	0.58
1997	2,471	10	0.40
1998	2,526	19	0.75
1999	2,815	20	0.71
2000	2,912	20	0.69
2001	2,985	41	1.37
2002	3,150	41	1.30
2003	3,434	37	1.08
2004	3,548	34	0.96
2005	3,487	39	1.12
2006	3,378	24	0.71
2007	3,406	26	0.76
2008	3,521	83	2.36
2009	3,700	55	1.49
2010	3,906	42	1.08
2011	3,904	39	1.00
2012	3,705	11	0.30

Table 3.2

Characteristics of the Final Sample: Breakdown by Size and Country

This table presents details on the characteristics of our final sample. Panel A shows descriptive statistics for the distribution of firms and firm-months across size classes. # of firms is the available number of firms for all years. # of firm-months is the number of observations. We provide also the relative fractions of total firms and firm-months that each size class represents. Finally, the column "Total MC" shows the total market capitalization of each size class averaged across the years of the study. We measure market capitalization in millions of euros. Panel B presents the breakdown of firms and firm-months by country, with corresponding percentages. Start date is the date at which the information on firms of a given country starts to be available; the end date in our sample, December 2012, is the same for all countries.

Panel A. Breakdown by Size

Segment	Size	# of firms	(%)	# of firm-months	(%)	Total MC	(%)
Nano cap	< 10 mio	1,419	16.81	106,570	13.22	7,401	0.11
Micro cap	< 50 mio	2,631	31.18	219,273	27.20	68,153	1.03
Small cap	< 150 mio	1,678	19.88	158,265	19.63	150,178	2.27
Mid cap	< 1 bio	1,855	21.98	205,855	25.54	735,025	11.11
Large cap	< 50 bio	839	9.94	112,526	13.96	4,239,777	64.07
Mega cap	≥ 50 bio	17	0.20	3,668	0.45	1,417,300	21.42
Overall sample		8,439		806,157		6,617,834	

Panel B. Breakdown by Country

Country	Start date	# of firms	(%)	# of firm-months	(%)
Austria	Jan-90	112	1.33	11,676	1.45
Belgium	Jan-90	151	1.79	17,842	2.21
Bulgaria	Mar-08	130	1.54	4,009	0.50
Czech Republic	Mar-98	71	0.84	3,679	0.46
Denmark	Jan-90	195	2.31	24,151	3.00
Finland	Jan-90	146	1.73	18,589	2.31
France	Jan-90	1,126	13.34	111,829	13.87
Germany	Jan-90	1,104	13.08	112,428	13.95
Greece	Oct-90	315	3.73	35,558	4.41
Hungary	Mar-95	45	0.53	3,558	0.44
Ireland	Jan-90	68	0.81	8,549	1.06
Italy	Jan-90	340	4.03	37,353	4.63
Netherlands	Jan-90	213	2.52	28,940	3.59
Norway	Jan-90	290	3.44	24,632	3.06
Poland	Mar-95	249	2.95	10,620	1.32
Portugal	Oct-90	94	1.11	10,002	1.24
Romania	Mar-02	65	0.77	2,690	0.33
Serbia	Jan-12	47	0.56	445	0.06
Spain	Jan-90	175	2.07	22,619	2.81
Sweden	Jan-90	525	6.22	42,856	5.32
Switzerland	Jan-90	232	2.75	31,695	3.93
United Kingdom	Jan-90	2,746	32.54	242,437	30.07
Overall Sample		8,439	100.00	806,157	100.00

Table 3.3

The Merton Measure and Other Firm Characteristics

The table presents descriptive statistics for the average Merton measure, monthly returns, size and BM ratio over the period January 1990 to December 2012. The sample spans 22 European countries. Monthly return is the time-series average of the cross-sectional average returns within each country. We measure returns in euros and express them in percent. Merton measure, size and BM are the time-series averages of the cross-sectional average Merton measures, market capitalization and BM ratios. We express the Merton measure in percentage terms (as it is a probability) and market capitalization in millions of euros.

	Merton measure			Monthly Returns			Size			BM		
Country	**Mean**	**Median**	**St. Dev.**	**Mean**	**Median**	**St. Dev.**	**Mean**	**Median**	**St. Dev.**	**Mean**	**Median**	**St. Dev.**
Austria	4.36	3.08	3.42	0.55	0.58	5.21	541.11	313.22	389.08	0.80	0.78	0.27
Belgium	4.70	3.96	2.88	0.63	0.90	4.16	963.20	889.92	492.07	0.84	0.81	0.18
Bulgaria	14.29	12.99	8.46	-0.64	0.28	8.18	33.02	25.84	22.61	1.74	1.83	0.35
Czech Republic	3.31	1.25	3.87	1.28	1.33	4.38	481.78	505.48	297.82	1.72	1.47	0.68
Denmark	4.09	2.76	3.10	0.69	0.78	4.64	580.10	489.44	337.62	0.90	0.93	0.23
Finland	4.11	2.63	4.63	0.95	0.52	6.22	1,247.22	1,129.59	848.93	0.74	0.69	0.25
France	5.00	4.28	2.53	0.77	0.94	4.63	1,557.86	1,619.24	576.00	0.82	0.81	0.18
Germany	4.67	3.76	3.07	0.55	0.79	3.93	1,457.07	1,443.94	431.23	0.70	0.64	0.23
Greece	6.71	4.61	5.79	1.01	-0.04	10.71	197.65	176.01	137.97	1.12	0.83	0.81
Hungary	9.14	8.76	5.24	1.62	1.15	9.39	82.84	83.77	39.55	1.33	1.33	0.48
Ireland	5.56	4.64	3.13	1.09	1.21	6.48	784.02	799.62	512.20	0.93	0.82	0.35
Italy	6.42	5.72	3.23	0.31	0.22	6.40	1,492.03	1,476.92	930.09	1.00	0.98	0.30
Netherlands	3.42	2.91	2.21	0.58	0.86	4.93	1,832.05	1,866.32	920.79	0.75	0.72	0.21
Norway	7.37	6.85	4.23	1.11	1.43	6.83	508.20	426.60	262.46	0.89	0.86	0.32
Poland	10.27	8.57	9.51	1.31	0.69	10.82	69.94	38.80	58.53	1.27	1.04	0.76
Portugal	7.31	6.69	4.04	0.85	0.20	5.70	659.05	635.25	453.15	1.15	1.11	0.30
Romania	16.69	13.03	10.09	2.02	1.33	9.11	87.34	39.06	87.19	2.15	2.12	0.53
Serbia	12.89	13.26	3.19	0.59	0.45	5.02	17.20	16.96	2.34	3.21	3.19	0.19
Spain	4.16	3.96	2.65	0.68	0.78	5.81	2,142.53	1,995.74	1,246.68	0.89	0.84	0.36
Sweden	6.64	6.21	4.13	1.02	0.91	7.05	1,084.76	886.09	681.68	0.77	0.74	0.28
Switzerland	3.13	2.34	2.41	0.75	0.95	4.54	2,187.33	2,356.76	961.16	0.88	0.83	0.26
United Kingdom	4.27	3.88	2.00	0.81	1.15	5.54	1,288.24	1,367.40	559.96	0.86	0.82	0.23
Overall Sample	5.84	4.44	5.10	0.86	0.82	6.50	1,006.23	765.26	896.78	0.99	0.86	0.50

Table 3.4

Portfolios sorted on the Physical PD

From January 1990 to December 2012, at the beginning of each month, we sort stocks into quintile portfolios based on their adjusted physical PD in the previous month. We adjust by dividing the physical PDs by the country average for this month. We report results with the Merton measure as a measure of physical PDs. Portfolio 1 is the portfolio with the lowest physical PD and portfolio 5 is the portfolio with the highest physical PD. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the equally and value-weighted portfolio returns and alphas. EMKT is the excess market return, SMB is the return difference between small stocks and big stocks, HML is the return difference between value stocks and growth stocks, and WML is the return difference between winning stocks and losing stocks. The column "High-Low" shows average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest physical PD stock quintile and short the lowest physical PD stock quintile. We denominate returns in euros and express them in percentage terms. Panel B shows loadings on the four factors from regressions of the equally and value-weighted excess returns. We calculate the t-statistics in parentheses from Newey-West (1987) standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel C reports PDs, size (in millions of euros), book-to-market ratios (BM), leverage ratios (LR) and return-on-assets (ROA) for each portfolio. SDR betas and IDR alphas are also reported and will be analyzed in further detail in the following tables.

Portfolios	High PD 5	4	3	2	Low PD 1	High-Low	*t*-stat
			Panel A. Portfolio Returns				
Equally-weighted							
Return	0.5776	0.5195	0.5985	0.6502	0.6436	-0.0660	(-0.25)
CAPM α	0.2379	0.1644	0.2569	0.3290	0.3453	-0.1075	(-0.44)
3-factor α	0.2269	0.1534	0.2386	0.3130	0.3296	-0.1027	(-0.48)
4-factor α	0.3575	0.2195	0.2922	0.3219	0.3197	0.0378	(0.16)
Value-weighted							
Return	0.2062	0.4758	0.4570	0.4597	0.6965	-0.4904	(-1.08)
CAPM α	-0.1955	0.0845	0.1014	0.1216	0.3982	-0.5936	(-1.40)
3-factor α	-0.2704	0.0450	0.1053	0.1518	0.4128	**-0.6832**	**(-1.84)***
4-factor α	-0.1777	0.1973	0.1828	0.1890	0.4675	**-0.6452**	**(-1.63)***
			Panel B. Four-Factor Regression Coefficients				
Equally-weighted							
EMKT	**0.238**	**0.245**	**0.205**	**0.167**	**0.133**		
	(3.00)**	**(3.66)****	**(3.35)****	**(3.08)****	**(3.12)****		
SMB	**1.036**	**0.961**	**0.848**	**0.699**	**0.524**		
	(6.51)**	**(6.60)****	**(7.00)****	**(6.57)****	**(6.25)****		
HML	0.121	0.132	0.143	0.134	0.115		
	(0.86)	(1.07)	(1.33)	(1.37)	(1.46)		
WML	-0.011	0.025	0.026	0.052	0.049		
	(-0.14)	(0.35)	(0.41)	(0.86)	(1.02)		

Value-weighted					
EMKT	**0.286**	**0.296**	**0.248**	**0.177**	**0.090**
	(3.28)**	**(3.56)****	**(2.91)****	**(2.52)****	**(2.09)****
SMB	**1.345**	**1.175**	**1.001**	**0.716**	**0.451**
	(6.79)**	**(6.65)****	**(6.69)****	**(5.55)****	**(5.24)****
HML	0.336	0.204	0.088	0.005	0.013
	(1.83)	(1.31)	(0.62)	-0.05	-0.15
WML	0.016	-0.034	0.008	0.035	0.021
	-0.14	(-0.35)	(0.09)	(0.43)	(0.31)
Panel C. Portfolio Characteristics					
Average PD	22.5600	1.7749	0.1614	0.0096	0.0000
Average Size	286.42	530.43	1,000.41	1,707.40	2,674.78
Average BM	1.4545	1.0046	0.7706	0.6097	0.4949
Average LR	4.0889	1.7436	1.0925	0.7103	0.4025
Average ROA	-0.0623	-0.0045	0.0177	0.0297	0.0369
Average SDR β	0.0590	0.1574	0.0770	0.0327	0.0060
Average IDR α	14.3208	0.7767	-0.0892	-0.0567	0.1510

Table 3.5

Summary Statistics on VIX

In this table, VIX is the CBOE volatility index and Δ_mVIX is the monthly change in VIX. Mean, Std, Skew, and Kurt refer to the mean, standard deviation, skewness, and kurtosis, respectively. VSTOXX, VFTSE and VDAX are the EUROSTOXX 50, FTSE 100 and DAX volatility indices, which follow the VIX methodology for the European, UK, and German markets respectively. Δ_mEurostoxx50 is the monthly change in EUROSTOXX 50 and Δ_mMSCIEurope is the monthly change in MSCI Europe. EMKT is the value-weighted excess return on the European market portfolio over the risk-free rate and SMB and HML are the Fama-French factors for Europe. Aggregate PD is the monthly average and Median PD is the monthly median of the Merton measure values of all firms. We calculate the t-statistics from Newey-West (1987) standard errors (up to five lags).

Panel A. Summary Statistics on VIX and VIX Monthly Changes (Δ_mVIX)				
	Mean	Std	Skew	Kurt
VIX	20.1978	8.0533	2.0133	10.1303
Δ_mVIX	-0.0267	4.2391	0.8229	8.1017
Panel B. Correlation between VIX and Other Volatility Indices				
	VSTOXX	VFTSE	VDAX	
VIX	0.9100	0.9449	0.9492	
Panel C. Correlation between Δ_mVIX and European Stock Indices				
	Δ_mEUROSTOXX50	Δ_mMSCIEurope		
Δ_mVIX	-0.6335	-0.5835		
Panel D. Correlation between Δ_mVIX and Other Factors				
	EMKT	SMB	HML	
Δ_mVIX	-0.1743	-0.1670	-0.0623	
Panel E. Time-Series Regression of the Aggregate and Median Merton measure on VIX				
	Constant	VIX	*R*-squared	
Aggregate PD	1.8060	0.1534	0.2686	
	(5.43)	(10.07)		
Median PD	-0.4676	0.0026	0.3112	
	(-8.30)	(11.17)		

Table 3.6

Portfolios sorted on the SDR Beta

From January 1992 to December 2012, at the beginning of each month, we sort stocks into quintile portfolios based on their adjusted SDR beta in the previous month. We adjust the SDR betas by dividing them by the country average for this month. We report results with the VIX SDR beta, which we measure as the coefficient (sensitivity) from 24-months rolling regressions of the PD on VIX. Portfolio 1 is the portfolio with the lowest SDR beta and portfolio 5 is the portfolio with the highest SDR beta. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the equally and value-weighted portfolio returns and alphas. EMKT is the excess market return, SMB is the return difference between small stocks and big stocks, HML is the return difference between value stocks and growth stocks, and WML is the return difference between winning stocks and losing stocks. The column "High-Low" shows average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest SDR beta stock quintile and short the lowest SDR beta stock quintile. We denominate returns in euros and express them in percentage terms. Panel B shows loadings on the four factors from regressions of the equally and value-weighted excess returns. We calculate the t-statistics in parentheses from Newey-West (1987) standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel C reports PDs, SDR betas, IDR alphas, sizes (in millions of euros), BM, LR and ROA ratios for each portfolio.

Portfolios	High β 5	4	3	2	Low β 1	High-Low	*t*-stat
			Panel A. Portfolio Returns				
Equally-weighted							
Return	0.8924	0.7232	0.7175	0.7041	0.5985	**0.2939**	**(1.80)***
CAPM α	0.5249	0.3922	0.3889	0.3777	0.2700	**0.2549**	**(1.80)***
3-factor α	0.4577	0.3070	0.3014	0.3317	0.1835	**0.2742**	**(1.89)***
4-factor α	0.4460	0.2883	0.2750	0.2697	0.1117	**0.3343**	**(1.97)****
Value-weighted							
Return	0.8066	0.6384	0.5877	0.5687	0.4391	0.3675	(1.24)
CAPM α	0.4162	0.3016	0.2814	0.2720	0.0859	0.3302	(1.14)
3-factor α	0.3149	0.3152	0.2153	0.2297	0.0985	0.2164	(0.76)
4-factor α	0.4035	0.3061	0.1989	0.1854	0.0527	0.3508	(1.19)
			Panel B. Four-Factor Regression Coefficients				
Equally-weighted							
EMKT	**0.266**	**0.182**	**0.182**	**0.191**	**0.191**		
	(3.63)**	**(3.16)****	**(3.72)****	**(3.53)****	**(3.71)****		
SMB	**0.979**	**0.715**	**0.767**	**0.726**	**0.771**		
	(6.28)**	**(5.66)****	**(7.02)****	**(6.67)****	**(6.30)****		
HML	0.148	0.197	**0.204***	0.118	**0.216**		
	(1.16)	(1.86)	**(2.05)****	(1.21)	**(2.16)****		
WML	0.009	0.015	0.021	0.050	0.058		
	(0.13)	(0.23)	(0.37)	(0.81)	(0.91)		

Value-weighted					
EMKT	**0.287**	**0.214**	**0.137**	**0.133**	**0.250**
	(3.60)**	**(3.82)****	**(2.54)****	**(2.45)****	**(2.38)****
SMB	**1.060**	**0.763**	**0.652**	**0.683**	**0.770**
	(5.68)**	**(5.04)****	**(6.28)****	**(6.81)****	**(4.55)****
HML	0.196	-0.040	0.152	0.104	-0.026
	(1.37)	(-0.36)	(1.47)	(1.08)	(-0.19)
WML	-0.071	0.007	0.013	0.036	0.037
	(-0.72)	(0.09)	(0.17)	(0.47)	(0.43)
Panel C. Portfolio Characteristics					
Average PD	10.7144	1.6788	0.5810	0.6172	8.7870
Average SDR β	0.8881	0.0516	0.0081	-0.0025	-0.6166
Average IDR α	-5.9819	0.3573	0.3122	0.5973	18.7048
Average size	708.81	1,691.08	1,957.37	1,964.43	1,044.72
Average BM	1.1773	0.7985	0.6703	0.6806	1.0280
Average LR	2.8791	1.2538	0.8418	0.8740	2.2675
Average ROA	-0.0251	0.0144	0.0272	0.0256	-0.0093

Table 3.7

Portfolios sorted on the IDR Alpha

From January 1992 to December 2012, at the beginning of each month, we sort stocks into quintile portfolios based on their adjusted IDR alpha in the previous month. We adjust the IDR alphas by dividing them by the country average for this month. We report results with the IDR alpha, which we measure as the constant from 24-months rolling regressions of the PD on VIX. Portfolio 1 is the portfolio with the lowest IDR alpha and portfolio 5 is the portfolio with the highest IDR alpha. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the equally and value-weighted portfolio returns and alphas. EMKT is the excess market return, SMB is the return difference between small stocks and big stocks, HML is the return difference between value stocks and growth stocks, and WML is the return difference between winning stocks and losing stocks. The column "High-Low" shows average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest IDR alpha stock quintile and short the lowest IDR alpha stock quintile. We denominate returns in euros and express them in percentage terms. Panel B shows loadings on the four factors from regressions of the equally and value-weighted excess returns. We calculate the t-statistics in parentheses from Newey-West (1987) standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel C reports PDs, SDR betas, IDR alphas, sizes (in millions of euros), BM, LR and ROA ratios for each portfolio.

Portfolios	High α 5	4	3	2	Low α 1	High-Low	*t*-stat
			Panel A. Portfolio Returns				
Equally-weighted							
Return	0.6686	0.5484	0.7229	0.8372	0.8545	-0.1858	(-1.17)
CAPM α	0.3437	0.2203	0.3940	0.5016	0.4904	-0.1467	(-1.04)
3-factor α	0.2648	0.1605	0.3193	0.4281	0.4049	-0.1401	(-0.93)
4-factor α	0.1907	0.0888	0.2933	0.3885	0.4263	-0.2357	(-1.32)
Value-weighted							
Return	0.4450	0.4243	0.5613	0.6981	0.9573	**-0.5124**	**(-1.97)****
CAPM α	0.0847	0.1073	0.2678	0.3691	0.5894	**-0.5046**	**(-2.00)****
3-factor α	0.0675	0.0905	0.2106	0.3434	0.5682	**-0.5007**	**(-1.86)***
4-factor α	0.0504	0.0209	0.2197	0.3058	0.6456	**-0.5952**	**(-1.81)***
			Panel B. Four-Factor Regression Coefficients				
Equally-weighted							
EMKT	**0.192**	**0.192**	**0.184**	**0.199**	**0.245**		
	(3.53)**	**(4.10)****	**(3.25)****	**(3.65)****	**(3.35)****		
SMB	**0.851**	**0.706**	**0.740**	**0.756**	**0.903**		
	(6.40)**	**(5.95)****	**(6.73)****	**(6.60)****	**(6.18)****		
HML	0.198	0.154	0.174	0.175	0.182		
	(1.93)	(1.68)	(1.69)	(1.69)	(1.42)		
WML	0.060	0.058	0.021	0.032	-0.017		
	(0.88)	(1.02)	(0.32)	(0.52)	(-0.24)		

Value-weighted					
EMKT	**0.265**	**0.174**	**0.110**	**0.196**	**0.260**
	(2.51)**	**(3.51)****	**(2.01)****	**(3.33)****	**(3.27)****
SMB	**0.975**	**0.608**	**0.622**	**0.738**	**0.978**
	(5.50)**	**(5.47)****	**(6.37)****	**(5.85)****	**(5.23)****
HML	0.032	0.053	0.123	0.062	0.011
	(0.24)	(0.52)	(1.22)	(0.61)	(0.08)
WML	0.014	0.056	-0.007	0.030	-0.062
	(0.13)	(0.79)	(-0.10)	(0.38)	(-0.61)
Panel C. Portfolio Characteristics					
Average PD	14.1359	0.9728	0.3755	0.9788	5.9189
Average SDR β	-0.5192	0.0186	0.0159	0.0586	0.7511
Average IDR α	22.5424	0.4605	0.0017	-0.3840	-8.5594
Average size	685.71	1,731.89	2,058.58	1,792.29	1,096.63
Average BM	1.2175	0.7378	0.6510	0.7336	1.0144
Average LR	3.3213	1.0551	0.7350	1.0685	1.9494
Average ROA	-0.0328	0.0195	0.0291	0.0203	-0.0036

Table 3.8

Portfolios sorted on the Physical PD controlled by the SDR beta

From January 1992 to December 2012, at the beginning of each month, we sort stocks into three portfolios based on their SDR beta in the previous month. Within each portfolio, we further sort the stocks into three portfolios, based on their past month's PD. We adjust both SDR betas and PDs by the country average for this month. The sequential two-sort procedure produces nine portfolios in total. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the value-weighted returns for the nine portfolios as well as average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest PD stock portfolio and short the lowest PD stock portfolio for all three SDR beta tertiles. We denominate returns in euros and express them in percentage terms. We calculate t-statistics in parentheses from Newey-West (1987) standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel B reports PDs, SDR betas, IDR alphas, sizes (in millions of euros), BM, LR and ROA ratios for each portfolio.

	High PD	Medium PD	Low PD	High-Low	*t*-stat
Panel A. Portfolio Returns					
Return					
High β	0.5417	0.5244	0.7706	-0.2289	(-0.52)
Medium β	0.4619	0.5882	0.7573	-0.2954	(-0.92)
Low β	0.3950	0.4577	0.5273	-0.1323	(-0.43)
CAPM α					
High β	0.1583	0.1259	0.4206	-0.2622	(-0.60)
Medium β	0.1106	0.2710	0.4666	-0.3560	(-1.19)
Low β	0.0219	0.1123	0.2382	-0.2163	(-0.74)
3-factor α					
High β	-0.0891	0.0599	0.4366	-0.5256	(-1.52)
Medium β	0.0302	0.2105	0.4493	-0.4191	(-1.63)
Low β	-0.1291	0.0557	0.2529	-0.3820	(-1.53)
4-factor α					
High β	-0.1206	0.1737	0.4320	-0.5525	(-1.56)
Medium β	0.1028	0.1532	0.4304	-0.3276	(-1.14)
Low β	-0.2094	-0.0754	0.2392	-0.4486	(-1.54)
Panel B. Portfolio Characteristics					
Average Probability of Default					
High β	19.1166	2.2405	0.3490		
Medium β	1.9067	0.0481	0.0016		
Low β	16.2516	0.3921	0.0012		
Average SDR Beta					
High β	1.0535	0.4216	0.2041		
Medium β	0.0176	0.0099	0.0023		
Low β	-0.8949	-0.1651	-0.0532		
Average IDR Alpha					
High β	-3.5759	-4.3234	-2.4638		
Medium β	1.1159	-0.0774	-0.0148		
Low β	28.3705	4.6544	1.4518		
Average Size					
High β	362.54	845.58	2,020.93		
Medium β	853.24	1,942.74	3,026.05		
Low β	396.84	1,126.23	2,683.99		

Average Book-to-Market			
High β	1.4525	0.9930	0.6714
Medium β	0.9187	0.6279	0.4913
Low β	1.3291	0.8114	0.5432
Average Leverage Ratio			
High β	4.2396	1.7275	0.8601
Medium β	1.4752	0.7502	0.4095
Low β	3.4413	1.2164	0.5498
Average Return-on-Assets			
High β	-0.0529	-0.0029	0.0236
Medium β	0.0089	0.0308	0.0378
Low β	-0.0393	0.0164	0.0349

Table 3.9

Portfolios sorted on the Physical PD controlled by the IDR alpha

From January 1992 to December 2012, at the beginning of each month, we sort stocks into three portfolios based on their IDR alpha in the previous month. Within each portfolio, we further sort the stocks into three portfolios, based on their past month's PD. We adjust both IDR alphas and PDs by the country average for this month. The sequential two-sort procedure produces nine portfolios in total. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the value-weighted returns for the nine portfolios as well as average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest PD stock portfolio and short the lowest PD stock portfolio for all three IDR alpha tertiles. We denominate returns in euros and express them in percentage terms. We calculate t-statistics in parentheses from Newey-West (1987) standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel B reports PDs, SDR betas, IDR alphas, sizes (in millions of euros), BM, LR and ROA ratios for each portfolio.

	High PD	Medium PD	Low PD	High-Low	*t*-stat
		Panel A. Portfolio Returns			
Return					
High α	-0.1105	0.2944	0.6686	**-0.7791**	**(-1.90)***
Medium α	0.6117	0.4217	0.6185	-0.0068	(-0.03)
Low α	0.8658	0.8218	0.8121	0.0537	(0.15)
CAPM α					
High α	-0.4474	-0.0749	0.3600	**-0.8074**	**(-2.03)****
Medium α	0.2514	0.1110	0.3369	-0.0855	(-0.34)
Low α	0.4839	0.4371	0.4807	0.0032	(0.01)
3-factor α					
High α	-0.5367	-0.2476	0.3854	**-0.9221**	**(-2.45)****
Medium α	0.1781	0.0852	0.3037	-0.1256	(-0.61)
Low α	0.3494	0.3890	0.4682	-0.1188	(-0.43)
4-factor α					
High α	-0.6182	-0.3495	0.3408	**-0.9591**	**(-2.36)****
Medium α	0.2728	0.0637	0.2931	-0.0204	(-0.08)
Low α	0.3286	0.4680	0.4682	-0.1397	(-0.51)
		Panel B. Portfolio Characteristics			
Average Probability of Default					
High α	24.7076	2.0970	0.0269		
Medium α	1.4230	0.0179	0.0002		
Low α	10.9104	0.9728	0.1867		
Average SDR Beta					
High α	-0.6487	-0.1936	-0.0710		
Medium α	0.0488	0.0079	0.0017		
Low α	1.0079	0.2979	0.1347		
Average IDR Alpha					
High α	30.4494	8.2601	2.6320		
Medium α	0.0790	-0.0522	-0.0042		
Low α	-10.3704	-3.8171	-1.8405		
Average Size					
High α	304.65	695.78	2,305.09		
Medium α	869.65	1,964.63	3,044.43		
Low α	525.11	1,205.18	2,337.58		

Average Book-to-Market			
High α	1.4969	0.9854	0.6060
Medium α	0.8975	0.6193	0.4949
Low α	1.2925	0.8516	0.5959
Average Leverage Ratio			
High α	4.7897	1.8378	0.7441
Medium α	1.3513	0.6996	0.3793
Low α	2.8937	1.3207	0.6827
Average Return-on-Assets			
High α	-0.0606	-0.0063	0.0274
Medium α	0.0104	0.0318	0.0392
Low α	-0.0303	0.0138	0.0315

Table 3.10

Portfolios sorted on the SDR beta controlled by the physical PD

From January 1992 to December 2012, at the beginning of each month, we sort stocks into three portfolios based on their PD in the previous month. Within each portfolio, we further sort the stocks into three portfolios, based on their past month's SDR beta. We adjust PDs and SDR betas by the country average for this month. The sequential two-sort procedure produces nine portfolios in total. The portfolios are held for one month and are then rebalanced. Panel A shows the time-series monthly average of the value-weighted returns, respectively, for the 9 portfolios as well as average monthly raw returns and alphas (excess risk-adjusted returns) for portfolios that are long the highest SDR beta stock portfolio and short the lowest SDR beta stock portfolio for all three PD tertiles. We denominate returns in euros and express them in percentage terms. We calculate t-statistics in parentheses from Newey-West (1987) heteroskedasticity and autocorrelation-consistent standard errors. ** denotes significance at a 5% level and * at a 10% level. Panel B reports PDs, SDR betas, IDR alphas, sizes (in millions of euros), BM, LR and ROA ratios for each portfolio.

	High β	Medium β	Low β	High-Low	*t*-stat
Panel A. Portfolio Returns					
Return					
High PD	0.4605	0.4083	0.4810	-0.0204	(-0.06)
Medium PD	0.5666	0.4726	0.5833	-0.0168	(-0.09)
Low PD	0.8064	0.6782	0.4111	**0.3953**	**(2.63)****
CAPM α					
High PD	0.0607	0.0277	0.1303	-0.0696	(-0.21)
Medium PD	0.1960	0.1172	0.2209	-0.0248	(-0.12)
Low PD	0.5002	0.3909	0.1221	**0.3781**	**(2.49)****
3-factor α					
High PD	-0.0488	-0.1139	-0.0246	-0.0242	(-0.07)
Medium PD	0.1393	0.0625	0.1728	-0.0334	(-0.16)
Low PD	0.5147	0.3431	0.1201	**0.3946**	**(2.35)****
4-factor α					
High PD	-0.0637	0.0353	-0.0020	-0.0616	(-0.19)
Medium PD	0.1528	0.0401	0.1838	-0.0310	(-0.14)
Low PD	0.4314	0.2986	0.0712	**0.3602**	**(1.94)***
Panel B. Portfolio Characteristics					
Average Probability of Default					
High PD	16.9449	6.3173	16.4642		
Medium PD	0.2626	0.1492	0.1530		
Low PD	0.0016	0.0012	0.0016		
Average SDR Beta					
High PD	1.1806	0.0673	-0.9207		
Medium PD	0.3340	0.0172	-0.1212		
Low PD	0.0717	0.0022	-0.0374		
Average IDR Alpha					
High PD	-6.5717	2.8973	28.8355		
Medium PD	-3.9248	-0.4568	3.6604		
Low PD	-0.8194	-0.0064	1.1237		
Average Size					
High PD	391.39	540.80	374.53		
Medium PD	1,205.64	1,300.03	1,138.04		
Low PD	2,703.03	2,878.01	2,725.56		

Average Book-to-Market			
High PD	1.3957	1.1714	1.3493
Medium PD	0.8010	0.7462	0.7814
Low PD	0.5462	0.5162	0.5311
Average Leverage Ratio			
High PD	3.9661	2.3535	3.4811
Medium PD	1.2367	1.0154	1.0950
Low PD	0.5651	0.4552	0.4941
Average Return-on-Assets			
High PD	-0.0476	-0.0175	-0.0401
Medium PD	0.0144	0.0227	0.0191
Low PD	0.0329	0.0373	0.0361

3.8 Figures of Chapter 3

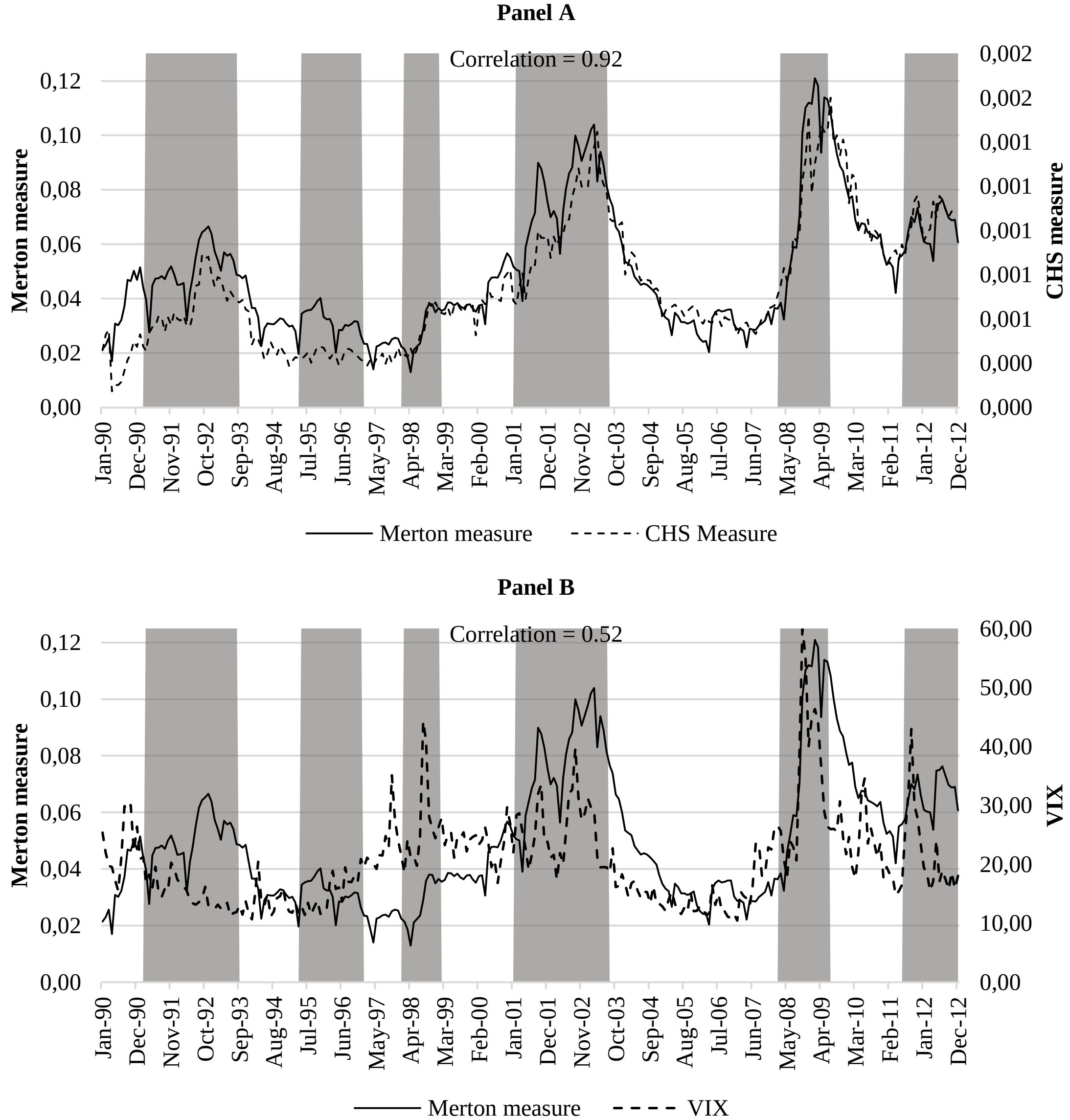

Figure 3.1. Merton Measure, Campbell, Hilscher and Szilagyi Measure and Volatility Index. The figure plots the monthly aggregate Merton (left scale) and CHS (right scale) measures for firms in the overall sample (Panel A) and the monthly aggregate Merton measure (left scale) and monthly VIX (right scale) values (Panel B). The columns denote recession periods in the euro area, as indicated by the OECD.

3.9 Appendices of Chapter 3

Appendix 3.1 The Merton Measure

Following Vassalou and Xing (2004), we allow only equity and debt in the capital structure of the firm. In Merton's model, equity can be viewed as a call option on the firm's assets with a strike price equal to the value of debt. The reason is that equity is a residual claim, i.e. equity holders lay claim to all the cash flows left over only after all the debt holders have been satisfied.

The market value of the firm's assets follow a geometric Brownian motion as below:

$$dV_A = \mu V_A dt + \sigma_A V_A dW, \tag{A.1}$$

where V_A is the market value of the firm's assets, with an instantaneous drift μ, and instantaneous volatility σ_A. W is a standard Wiener process.

The market value of the firm's equity is given by the Black and Scholes (1973) formula for call options:

$$V_E = V_A N(d_1) - Xe^{-rT} N(d_2), \tag{A.2}$$

$$d_1 = \frac{\ln\left(\frac{V_A}{X}\right) + \left(r + \frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}},\ d_2 = d_1 - \sigma_A\sqrt{T}, \tag{A.3}$$

where V_E is the market value of firm's equity, X is the book value of debt that has a maturity equal to T, r is the risk-free rate, and N is the cumulative density function of the standard normal distribution.

First, we calculate the volatility of equity σ_E from the daily data of the past 12 months and use it as the initial value for the estimation of σ_A. Then, from (A.2) and (A.3), we compute V_A for each trading day of the past 12 months using V_E of that day and X. As Vassalou and Xing (2004) and KMV do, we use current liabilities (WC03101) plus half the long-term debt (WC03251) to calculate the book value of debt X. Also, to account for reporting delays that may influence data availability, we use the book value of debt at the fiscal year end, only after 4 months have passed from the fiscal year end. From the daily values of V_A we calculate σ_A for the next iteration. We repeat this process until the values of σ_A from two consecutive observations converge. Once we obtain a converged value of σ_A, we use it to find V_A from (A.2) and (A.3). We repeat the process at the end of every month and obtain monthly values for σ_A. We use the 1-year T-bill rate at the end of the month as the risk-free rate. Once we obtain daily values for V_A, we compute the drift μ as the mean of the change in $\ln V_A$. Finally, using the normal distribution implied by Merton, we can show that the physical PD at time t is given by the following formula:

$$PD_t = N(-Distance\ to\ Default_t) = N(-\frac{\ln\left(\frac{V_{A,t}}{X_t}\right)+\left(\mu-\frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}}), \tag{A.4}$$

where we refer to PD_t as the Merton measure.

In order to evaluate the performance of the Merton measure, we employ two widely used measures, the Hosmer and Lemeshow (2010) grouping and the area under the Receiver Operating Characteristic (ROC) curve.

First, based on the Hosmer and Lemeshow method, each month we rank the estimated PDs and divide them into deciles. Out of the ten groups created (each one containing 1/10 of the firms in that month), the first group has the smallest average estimated PD and the last the largest. Next, we aggregate the number of defaulted firms in each decile for each month over the sample period and calculate the corresponding percentages of the defaulted firms in each decile. The percentage of defaulted firms in the last decile is 58.72%. When we look at the last three deciles, this percentage becomes 79.87%. This provides us initial evidence that the Merton measure captures important default-related information.

Second, we construct the area under the ROC curve (AUC) from the estimated PDs versus the actual status of the firms in each month for all possible cut-off probability values. Specifically, the curve plots the ratio of correctly classified defaulted firms to actual defaulted firms (sensitivity) and the ratio of wrongly classified healthy firms to actual healthy firms (1 - specificity) for all possible cut-offs. The AUC ranges from zero to one. A model with an AUC close to 0.5 is considered a random model with no discriminatory power. An AUC of 0.7 to 0.8 represents good discriminatory power, an AUC of 0.8 to 0.9 very good discriminatory power and an AUC over 0.9 is exceptional and extremely unusual. The AUC that we obtain is equal to 0.8212. This result further supports our belief that the Merton measure is indeed a good default predictor.

As a supplementary and final test, we follow Vassalou and Xing (2004) and compare the PDs of the defaulted firms (treatment group) with the PDs of a group of non-defaulted firms (control group). For each defaulted firm, we choose a healthy firm of similar size (market capitalization) and same industry (4-digit ICB code). We try to match the size of defaulted and healthy firms on the exact month or year of delisting due to default whenever possible. Figure 3.1.1 shows the average PDs of both groups up to 160 months before delisting. It is apparent that the PDs of both groups move closely together up to four years (48 months) before delisting. In the beginning of the fourth year before delisting though, the average PD of the treatment group goes up sharply, whereas the average PD of the control group does not follow this extreme behavior. Its moderate upward movement can be attributed to general worsening economic conditions in times of many defaults that move upward all PDs in the economy. The average PD at $t = 0$ is 0.14 for healthy firms and 0.34 for defaulted firms (around 2.5 times higher). This final test provides additional support that the Merton measure captures default risk successfully.

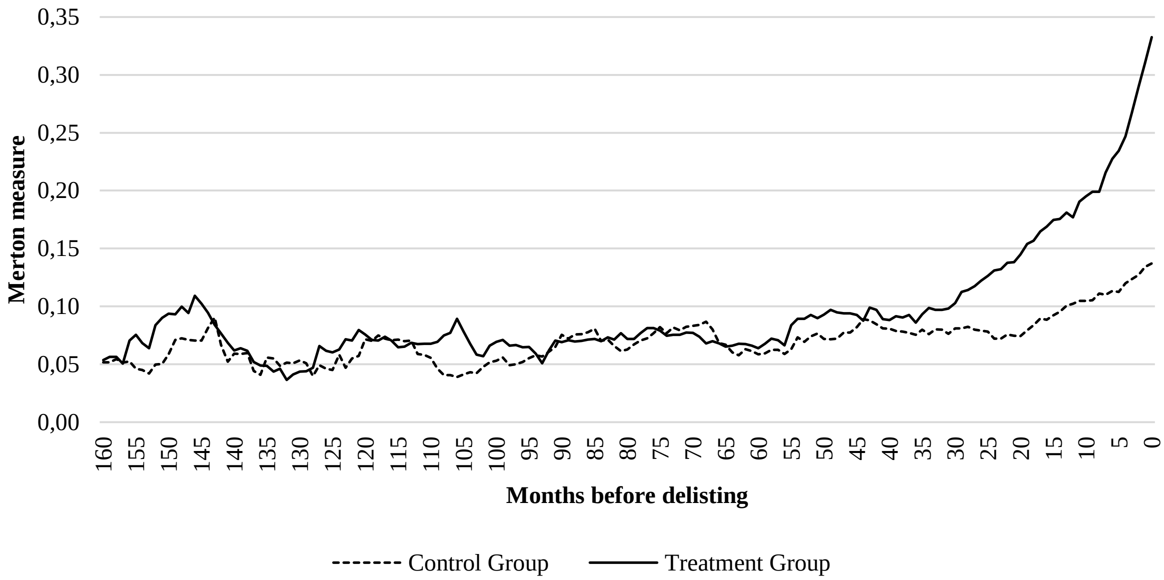

Figure 3.1.1. Average Merton Measure of treatment group (defaulted firms) and control group (healthy firms). We choose firms in the control group that have similar size (market capitalization) and the same four-digit industry code as those in the treatment group. Specifically for size, we select firms that have similar size with their defaulted counterparts immediately before they delist. We also match the month or year of delisting whenever possible.

Appendix 3.2 Calculation of the CHS Measure

Following Campbell, Hilscher and Szilagyi (2008), we use eight variables to calculate the CHS measure (all converted into euros). NIMTA is the ratio of net income (WC07250) to the market-adjusted version of total assets, where the latter is the sum of the market value of equity and the book value of liabilities (WC03351); TLMTA is the ratio of total liabilities to the market-adjusted version of total assets; EXRET is the monthly log excess return relative to the MSCI index of the country that is the firm's main market; SIGMA is the standard deviation of the daily returns over the previous year; RSIZE is the log ratio of firm's market value to the total market value of firms in the same country and month; CASHMTA is the ratio of cash and short-term investments (WC02001) to the market-adjusted version of total assets; MB is the market-to-book ratio; and PRICE is the log price per share winsorized at the first and third quartiles of the pooled price distribution. We winsorize all other variables at the first and ninety-ninth percentile of their pooled distributions. We lag all accounting data by at least 4 months and market data by 1 month, to ensure their availability at the time of default prediction. To avoid excluding firms shortly before they default, we use data for up to 12 months if more recent data are unavailable.

Table 3.2.1 presents summary statistics of these variables. A comparison of Panels B and C reveals the differences in the defaulted observations. They have lower profitability, higher leverage, lower stock excess returns, higher stock volatility, lower MB ratios and lower prices compared to the healthy observations. They are also smaller. Interestingly, they do not differ much in their cash holdings.

Concerning the applied estimation method, we assume that the marginal probability of default over the next period follows a logistic distribution and is given by:

$$PD(Y_{i,t} = 1|x_{i,t-1}) = \frac{1}{1 + \exp(-\alpha - \beta x_{i,t-1})}, \tag{B.1}$$

where $Y_{i,t}$ is an indicator that equals one if the firm defaults in period t and zero otherwise, i.e. if the firm disappears from the sample for some reason other than default, such as delisting due to a merger; and $\beta x_{i,t-1}$ is a function of firm-specific characteristics that includes a vector of predictor variables $x_{i,t-1}$ known at the end of the previous period. Finally, to capture cross-country differences, we follow two different methods: (i) we estimate separate models for each country; (ii) we introduce country fixed effects and estimate only one model.

Table 3.2.2 reports the regression results only under method (ii) due to space limitations. The coefficients confirm the findings from Table 3.2.1. The CHS measure is negatively related to profitability (NIMTA), excess return (EXRET), size (RSIZE), and PRICE. It is positively related to leverage (TLMTA), volatility (SIGMA), liquidity (CASHMTA) and MB. Most coefficients are significant at a 5% level, with the exception of CASHMTA and MB. The pseudo-R^2 (McFadden's R^2) is 17.4%, indicating a rather good fit. The pseudo-R^2 may look low when compared to R^2 values of linear regression models, but such low values are normal in logistic regression.

Table 3.2.1

Summary Statistics for the CHS Measure

The table reports summary statistics for all of the accounting and market variables used to calculate the CHS measure. NIMTA is the ratio of net income (WC07250) to the market-adjusted version of total assets, where the latter is the sum of the market value of equity and the book value of liabilities (WC03351); TLMTA is the ratio of total liabilities to the market-adjusted version of total assets; EXRET is the monthly log excess return relative to the MSCI index of the country that is the firm's main market; RSIZE is the log ratio of firm's market value to the total market value of firms in the same country and month; SIGMA is the standard deviation of the daily returns over the previous year; CASHMTA is the ratio of cash and short-term investments (WC02001) to the market-adjusted version of total assets; MB is the market-to-book ratio; and PRICE is the log price per share winsorized at the first and third quartiles of the pooled price distribution. All other variables are truncated at the first and ninety-ninth percentile of their pooled distributions. Panel A describes the distributions of the variables in all observations, Panel B describes the sample of healthy observations, and Panel C describes the defaulted observations.

	NIMTA	TLMTA	EXRET	RSIZE	SIGMA	CASHMTA	MB	PRICE
				Panel A. All				
Mean	0.01	0.45	0.00	-7.98	0.41	0.09	2.33	1.65
Median	0.03	0.44	-0.01	-8.04	0.36	0.06	1.63	1.62
Std.Dev.	0.06	0.23	0.10	2.41	0.20	0.09	2.03	1.03
Min	-0.16	0.07	-0.19	-12.01	0.15	0.00	0.40	0.39
Max	0.09	0.84	0.19	-3.61	0.91	0.33	8.22	2.92
N	761,779	761,897	796,573	803,106	803,106	761,578	802,965	803,106
				Panel B. Healthy				
Mean	0.01	0.45	0.00	-7.98	0.41	0.09	2.33	1.65
Median	0.03	0.44	-0.01	-8.04	0.36	0.06	1.63	1.62
Std.Dev.	0.06	0.23	0.10	2.41	0.20	0.09	2.03	1.03
N	761,257	761,374	795,979	802,511	802,511	761,055	802,370	802,511
				Panel C. Defaulted				
Mean	-0.07	0.64	-0.05	-10.56	0.66	0.10	1.48	0.74
Median	-0.08	0.77	-0.04	-11.39	0.71	0.05	0.69	0.39
Std.Dev.	0.09	0.25	0.13	1.88	0.24	0.11	1.90	0.74
N	522	523	594	595	595	523	595	595

Table 3.2.2

Regression Results for the CHS Measure

The table reports results from the multi-period logit regression of the default indicator on the eight predictor variables. NIMTA is the ratio of net income (WC07250) to the market-adjusted version of total assets, where the latter is the sum of the market value of equity and the book value of liabilities (WC03351); TLMTA is the ratio of total liabilities to the market-adjusted version of total assets; EXRET is the monthly log excess return relative to the MSCI index of the country that is the firm's main market; RSIZE is the log ratio of firm's market value to the total market value of firms in the same country and month; SIGMA is the standard deviation of the daily returns over the previous year; CASHMTA is the ratio of cash and short-term investments (WC02001) to the market-adjusted version of total assets; MB is the market-to-book ratio; and PRICE is the log price truncated at the first and third quartiles of the pooled price distribution. We truncate all other variables at the first and ninety-ninth percentile of their pooled distributions. We lag all accounting data by at least 4 months and market data by 1 month. The model is estimated for January 1990 to December 2012, with yearly observations. Parameter estimates are given first followed by chi-square values in parentheses. Standard errors are cluster-robust to correct for dependence between firm-year observations of the same firm. Numbers significant at the 5% level are in bold.

NIMTA	**-4.449**	**(-7.15)**
TLMTA	**2.914**	**(11.33)**
EXRET	**-1.550**	**(-3.91)**
RSIZE	**-0.455**	**(-10.29)**
SIGMA	**2.311**	**(9.19)**
CASHMTA	0.367	(0.75)
MB	0.014	(0.51)
PRICE	**-0.253**	**(-3.32)**
Constant	**-14.160**	**(-27.54)**
Firm-year observations	755,243	
Firms	7,980	
Distressed firms	522	
Country fixed effects	Yes	
Pseudo R-squared	0.174	
Log likelihood	-3568.9	
Wald test	970.0	

Chapter 4 Are Private Equity Backed Initial Public Offerings Any Different? Timing, Information Asymmetry and Post-IPO Survival

"I am not against Private Equity in general, but when it comes to IPOs they are in the business to get the highest price for their investors. This means there is a tendency to flatter the books to make the investment look a lot better than it is." (James Laing, Aberdeen Asset Management, in the *Financial Times*, 18 February 2014).

4.1 Introduction

In this study with the term Private Equity (PE) I refer to both buyouts (BOs) and venture capital (VC) transactions, which are the largest and most important subclasses of PE. BO and VC sponsors are similar in terms of involvement and contribution to their portfolio companies. The main differences between the two are the companies they invest in and the methods they use to create value. VC targets early stage companies with high growth potential (often start-ups based on new technology or other innovation) and uses minority equity investment. BOs target larger and more mature companies (typically with above average profit margins, tangible assets and stable cash flows) and often use leverage to finance acquisitions (Metrick and Yasuda, 2011). Usually studies focus on either BOs or VC, but here I examine both, differentiating between the two. I focus on the period just before, during and after exit of PE sponsors and examine only exits via IPOs.

There are three ways that a PE sponsor can exit a portfolio company: (i) a sale to another financial buyer (e.g. another PE fund - in thecase of BOs, this is a secondary BO), (ii) a sale to a strategic buyer (i.e. a trade sale), and (iii) an IPO. I study this latter IPO exit because the literature argues that they can potentially involve more information asymmetry than other exit strategies. Bayar and Chemmanur (2011) build a theoretical model that predicts that, in high IPO valuation periods, companies that are harder to value by public investors are more likely to go public than be acquired.[16] In another theoretical paper, Chemmanur and Fulghieri (1999) show that public investors produce less information than acquirers due to the free-rider problem.[17] Both studies suggest that information asymmetries between insiders and public investors can result in inflated valuations, but that acquirers (both financial and strategic ones) can value firms more accurately. This is because financial buyers perform sophisticated analyses and strategic buyers thoroughly investigate potential synergies before investing.

The role of PE sponsors, as professional insiders, in such a setting is interesting. On the one hand, they may be more able to "exploit" the IPO market than insiders of stand-alone companies. On the other, every effort to "fool" public investors may be detrimental for their reputation and, as a consequence, their liquidity. It is possible also that BO and VC sponsors behave differently from each other. Academic literature as well as practitioners suggest that they may have different motives when taking portfolio companies public. BO sponsors prefer to exit quickly as these deals are very large and involve high potential losses. Thus, they may rush companies into premature IPOs. In accordance with this argument, Cao (2011) finds that BO deals that are exited quickly default more often post-IPO. VC sponsors, instead, undertake more risk that BO sponsors as only a small percentage of their companies make it to an IPO (the "stars"). It is a real opportunity for VC sponsors to establish their reputation from these transactions and, as a

[16] In a following paper, Bayar and Chemmanur (2012) empirically confirm these findings.

[17] Whereas the information production cost is incurred only by a small group of public investors, the benefits are shared among all, reducing the incentive of any single public investor to engage in information production.

consequence, they are unlikely to "fool" the IPO market. In accordance with this intuition, Neus and Waltz (2005) show that VC sponsors have incentives to report the true quality of their portfolio companies during the IPO process.

In my study, I focus on IPO timing, information asymmetry and post-IPO survival. I also ask whether PE sponsors time their IPOs better, whether they inflate valuations more and whether they seek to sell firms with poor prospects ("unload lemons") compared to insiders of stand-alone companies. In my analysis I differentiate between BO and VC-backed IPOs in order to uncover their potentially different motives. My study is related to four literature strands that I present below.

The first is the strand on IPO timing, which refers to either performance or market timing. In the performance timing literature, studies test whether companies time their IPOs to occur in years when there are exceptionally favorable market fundamentals. In an early study, Degeorge and Zeckhauser (1993) study a sample of 62 reverse leveraged buyouts (RLBOs)[18] and find that LBO sponsors time their IPOs in years when their operating performance increases more than that of comparable companies. More recently, Cao (2011) studies 594 RLBOs from 1981 to 2006 and does not find evidence of performance timing. He does not include comparisons with other companies, unlike Degeorge and Zeckhauser (1993) which do make these comparisons. In the market timing literature, Ritter and Welch (2002) argue that market conditions are the most important factor in a company's decision to go public. Schulz (2003) characterizes this phenomenon as "pseudo" market timing because companies do not predict market peaks but simply follow their peers and go public at high valuation periods. Similarly, Alti (2005) shows theoretically that high offer price realizations have spillover effects that attract subsequent IPOs.

[18] These are LBO-backed IPOs.

Pastor and Veronesi (2005) develop similar predictions and confirm them empirically. For LBO sponsors, Cao (2011) studies the duration of LBO-backing pre-IPO and documents a negative relationship with high valuation periods and post-IPO default rates and a positive relationship with operating improvements. For VC sponsors, Lerner (1994) explores the exit choice between IPOs versus acquisitions and shows that they are particularly proficient at taking their portfolio companies public near market peaks. Similar are the findings of Ball, Chiu and Smith (2011) who argue in support of "pseudo" market timing. As Cao (2011) though, these two studies do not include comparisons with other firms.

The second strand is on IPO underpricing. Existing academic literature uses short-run underpricing (first-day returns) as a measure of information asymmetry and finds higher underpricing for VC-backed IPOs and lower for BO-backed IPOs.[19] Gompers (1996) argues that younger VC sponsors need to establish reputation in order to successfully raise capital for new funds (grandstanding hypothesis) and they use underpricing as a device to achieve this. For example, they might purposely leave money on the table to signal quality. Lee and Wahal (2004) report similar results with Gompers (1996) and find greater underpricing for VC-backed IPOs compared to other matched IPOs. Hogan, Olson and Kish (2001) find that first-day returns of RLBOs are higher from other IPOs in the period 1987-1998. However, these studies do not look at IPO proceeds to address the issue of inflated valuations.

The third strand focuses on default risk in PE transactions. These studies usually track companies only during the period that they are PE-backed, thus they examine default as an exit

[19] An exception is Megginson and Weiss (1991) who examine an early small sample of VC-backed IPOs and find that they are less underpriced than matched IPOs (certification role of VC). Barry et al. (1990) find lower underpricing for VC-backed IPOs with better monitors (monitoring role of VC).

outcome. On the BO side, two early studies of the first wave of BOs are Kaplan and Stein (1993) and Andrade and Kaplan (1998). Studies of the second wave of BOs often involve comparisons with other companies. Tykvova and Borell (2012), Hotchkiss, Smith, and Stromberg (2013) and Wilson and Wright (2013) document similar default rates between BO-backed and comparable companies in Europe, the US and the UK respectively. On the VC side, Puri and Zarutskie (2012) find that cumulative failure rates of VC-backed companies are lower relative to comparable companies. Of the above studies, only Hotchkiss, Smith, and Stromberg (2013) look also at defaults post-exit and find that BO-exited companies have lower default rates than others. This suggests that BO sponsors leave their portfolio companies in good financial condition. However, they do not study IPO exits in particular. I am aware of two more studies that examine default risk post-exit: Harford and Kolasinski (2013) for BO-backed companies and Jain and Kini (2000) for VC-backed companies. Harford and Kolasinski (2013) study strategic acquisitions and find that having been owned previously by a BO fund has no impact on whether the company will eventually undergo distressed restructuring. Jain and Kini (2000) study a small sample of 877 IPOs that took place in the period 1977-1990 and find that VC-backing increases the survival likelihood.

Whereas only Jain and Kini (2000) focus on post-IPO defaults, the fourth strand includes plenty of studies that focus on post-IPO stock and operating performance. On stock returns, Cao and Lerner (2009) and Brav and Gompers (1997) do not find strong evidence that BO and VC-backed IPOs respectively outperform others. On operating performance, early studies of a few RLBOs have apparently contradictory conclusions. Whereas Muscatella and Vetsuypens (1990) report improvements, Degeorge and Zeckhauser (1993) document declines post-IPO. Krishnan et al. (2011) study VC-backed IPOs and find that more reputable. VCs contribute to stronger post-

IPO operating performance. In a comparative framework, Holthausen and Larcker (1996) find that the operating performance of RLBOs is stronger than their industries as a whole.

I make several contributions to this literature. To my knowledge, this is the first study about IPO exits of both BO and VC-backed companies. By focusing on an environment of high information asymmetry and looking at both types of PE sponsors, I am better able to examine their behavior during the IPO process and also uncover potentially different incentives between them. Another interesting aspect is my focus on default risk (both pre and post-IPO) instead of operating performance. Although the two are related, default risk can be heavily influenced by other factors, such as leverage and liquidity. Thus default measures reflect better the company's pre-IPO financial situation. By looking also at actual defaults post-IPO, I offer important insights about the solvency situation of portfolio companies after PE sponsors unload them in the IPO market. It is also the first study to test if PE sponsors time their IPOs in "hot" periods or rush them more than their peers do. Finally, I extend the literature on underpricing and look directly at proceeds instead of only first-day returns, so enabling me to compare valuations of similar companies.

In the first part of my analysis, I examine the pre-IPO period in order to address IPO timing considerations for PE-backed versus stand-alone IPOs. Initially, I measure pre-IPO default risk instead of operating performance (Degeorge and Zeckhauser, 1993; Cao, 2011) and perform comparisons between PE-backed and stand-alone companies in a matching framework. After I control for leverage (in the case of BO-backing) and profitability (in the case of VC-backing) I do not find significant differences in the pre-IPO default measures of PE-backed and matched stand-alone IPOs. I interpret this as evidence that PE sponsors do not time their IPOs in years that their financial situation is better compared to their peers. Then, on the IPO market timing side, I move to a regression framework to test whether PE sponsors are better able than managers of stand-alone companies to time their IPOs for when overall market conditions are more favorable. I find neither BO nor VC-backed IPOs to be more common in hot periods. Moreover, although I find that, on average, companies enter at a younger age when the IPO market is hot, both BO and VC-backed IPOs that take place in hot market periods are older at the time of the IPO. The above suggests that PE sponsors do not target their IPOs in hot market periods and do not rush their companies into premature IPOs when market conditions are favorable more than managers of stand-alone companies do. If anything, they seem to rush them less.

In the next part of my analysis, I test whether information asymmetries during the IPO enable PE sponsors to inflate valuations compared to those achieved by stand-alone companies. Conversely, I investigate whether these sponsors' reputation and liquidity concerns mitigate such behavior. Consistent with the literature, I find more underpricing for VC-backed IPOs than stand-alone IPOs. However, I do not find significant differences in the first-day returns between BO-backed and stand-alone IPOs. As a next step, I go beyond past studies and look directly at the size of the IPO. Practitioners argue that the size of the IPO does not receive as much extensive attention in the media as first-day returns. By comparing the proceeds of BO and VC-backed IPOs with these of matched stand-alone IPOs, controlling for the float percentage, I can further test if these sponsors raise more or less cash. In the case of VC sponsors, the underpricing mentioned above together with past literature (Gompers, 1996; Neus and Walz, 2005; Lee and Wahal, 2004) suggests that they do not issue overvalued equity. However, the behavior of BO sponsors is less clear. As expected, I do not observe significant differences in IPO proceeds between VC-backed and matched stand-alone IPOs. Surprisingly, I find that BO-backed IPOs have lower IPO proceeds compared to matched stand-alone IPOs. These results suggest that BO sponsors not only do not inflate valuations, but on the contrary, public investors price BO-backed IPOs more conservatively. Thus these results contradict the criticism that BO sponsors have a tendency to inflate valuations.

In the last part of my analysis, I examine the post-IPO period in order to address the ultimate question of whether PE-backed IPOs delist more due to default or failure than stand-alone IPOs. I use actual default and failure rates post-IPO instead of stock or operating performance. With this I test whether PE sponsors are more likely than managers of stand-alone companies to take problematic portfolio companies public before hidden problems can unfold. By doing so they would transfer the risk and loss to public investors. Interestingly, I find that default and failure rates of BO-backed versus stand-alone IPOs and VC-backed versus stand-alone IPOs do not differ significantly in a matching framework that tracks companies up to five years post-IPO. When I move to a regression framework that uses all firm-year observations post-IPO, I even find evidence that BO and VC-backed IPOs default less frequently than others. Finally, in accordance with the theoretical model of Yung, Colak and Wang (2008), I find that IPOs that take place in hot market periods are significantly more likely to default, but this result is not any stronger for BO or VC-backed examples. These results indicate that PE sponsors are not any more likely than managers of stand-alone companies to "fool" the market and it is in accordance with the intuition that if these sponsors are caught "cheating", they will struggle to raise money in the future.

The paper is organized as follows: Section 4.2 describes the dataset and presents summary statistics. Section 4.3 describes the methodology and the reasons for its selection. Section 4.4 presents the results, and Section 4.5 concludes.

4.2 The Data

I collect data related to the IPOs that took place on the AMEX, NASDAQ and NYSE exchanges for the period 1975 to 2013. Data is provided by Jay Ritter [20] and also that which is available in Thomson Reuter's SDC Platinum New Issues Database. I follow the literature and exclude IPOs with an offer price below $5.00 per share, a total valuation below $1.5 million, unit offers, American Depository Receipts (ADRs), closed-end funds, natural resource partnerships, acquisition companies, Real Estate Investment Trusts (REITs), bank and Savings and Loans (S&L) IPOs, and firms not listed on CRSP. I end up with 7,033 IPOs out of which 897 are BO-backed (13%) and 2,763 are VC-backed (39%). I identify the BO and VC-backed IPOs as follows: For years 1984 to 2006, Jerry Cao provides me with his data on BO-backed IPOs. For years 2002 to 2013, Jay Ritter provides me with his data on VC-backed IPOs. For the remaining years, I complement these data with Thomson Reuter's VentureXpert Database. I follow Jay Ritter's classification and characterize growth-capital backed IPOs as VC-backed.[21]

I take offer dates, offer prices, file price ranges, proceeds, number of shares, SIC codes, headquarter states, over-allotment details and other IPO specific data items mostly from SDC Platinum New Issues Database and complement them where possible with data from Jay Ritter. I collect underwriter rankings and founding dates from Jay Ritter's website (based on Loughran and Ritter, 2004). To construct "IPO hotness" measures and to identify IPOs that receive star analyst coverage, I use the same website. Specifically, for my IPO hotness measures, I collect monthly data on the number of IPOs, average first-day returns and the percentage of deals each month that are priced above the midpoint of the original file price range (based on Ibbotson,

[20] Field and Karpoff (2002) and Loughran and Ritter (2004) use earlier versions of this dataset.

[21] Growth capital is a hybrid form between venture capital and buyouts, but closer to the first.

Sindelar, and Ritter, 1994). For IPOs that receive star coverage, I collect data on analysts (based on Cliff and Denis, 2004; Bradley and Ritter, 2008; Fang and Yasuda, 2009; Liu and Ritter, 2011). To calculate financial ratios, default measures and first-day returns, I get financial information from Compustat, and delisting and price information from CRSP. Compustat generally has financial information available up to two years pre-IPO. Codes DLDTE provide the delisting date and DLRSN the delisting reason. The standard matching variable is the 9-digit CUSIP, complemented with the ticker symbol where necessary.

Table 4.1 and Figure 4.1 provide summary statistics of my sample. Panel A of Table 4.1 shows the industry distribution of all IPOs, BO-backed IPOs and VC-backed IPOs across two-digit SIC codes that have at least 30 IPOs. There is significant variation across industries with BO-backed IPOs representing 45.10% of all IPOs in SIC code 53 (Food stores) and VC-backed IPOs representing the vast majority (64.38%) of all IPOs in SIC code 28 (Chemicals and Allied Products). In absolute terms, there are large clusters of both BO-backed and VC-backed IPOs in SIC code 73 (Business services). In Panel B, I show the geographic distribution of all IPOs, BO-backed IPOs and VC-backed IPOs noting states in which the company is headquartered. I display the distribution only for states that represent at least 2% of all IPOs (at least 140 IPOs). VC-backed IPOs represent 63.62% of all IPOs in Massachusetts and 59.80% in California. The most BO-backed IPOs are in California (130) and New York (77), and the most VC-backed IPOs are in California (1,694) and Massachusetts (306). There is high geographic concentration, with 45% of all IPOs originating from just four states: California, New York, Texas and Massachusetts. The fact that BO and VC sponsors invest in particular types of companies is reflected in both company and IPO characteristics, as shown in Panel C. BO-backed IPOs occur with older and larger companies (both in terms of assets and sales), whereas the opposite is the case for VC-backed IPOs. BO-backed IPOs also have higher valuations, gross spreads and underwriter rankings compared to both stand-alone and VC-backed IPOs. Moreover, they are more leveraged. VC-backed IPOs are less profitable on average. Past studies report very similar summary statistics (Lee and Wahal, 2004 and others). Consistent with the literature, VC-backed IPOs have higher first-day returns (27.86%) and BO-backed IPOs lower first-day returns (10.51%) than stand-alone IPOs. Finally, Figure 4.1 plots the number of all IPOs, BO-backed IPOs and VC-backed IPOs per year. It is obvious from the graph that both BO and VC-backed companies follow the trends of general IPO activity.

4.3 The Methodology

PE backing represents an endogenous choice for PE sponsors and entrepreneurs since its provision and receipt is the outcome of many negotiations between them (Megginson and Weiss, 1991). This endogenous choice is also reflected at the time of exit in the non-random distributions and characteristics of BO and VC-backed IPOs that Table 4.1 demonstrates. The non-randomness of these data suggests that I can use this information to construct instruments wiht some power to predict BO and VC backing. Once I construct such instruments that are correlated with the endogenous choice and control for the selection bias, I use them in a first-stage regression that predicts BO backing. Then, estimates from the first-stage regressions are fed into various matching methods to match BO-backed IPOs with stand-alone IPOs. I repeat exactly the same procedure for VC-backed IPOs. It is important to have two sets of estimates, one for BO and one for VC backing, since, as shown in Table 4.1, the two kinds of PE invest in quite different companies. In many cases, as an alternative to the above matching approach, I also estimate OLS regressions with two dummy variables, one for BO and one for VC backing, and various controls.

I now explain the matching approach in the case of BO backing (since the VC backing case is exactly the same). Let us suppose that t is a binary random variable that takes the value of 1 if an IPO is BO-backed and 0 otherwise. I call t the treatment level. Thus IPOs that are BO-backed belong to the treatment group and stand-alone IPOs belong to the control group. The potential outcome for a firm would when given treatment $t \in \{0,1\}$ is y_t. Each y_t has realizations y_{ti}. For potential outcome variables I examine pre-IPO default measures, underpricing, IPO proceeds and post-IPO delistings. The parameter of interest is the average treatment effect on the treated (ATET) which is the average effect of BO-backing on BO-backed IPOs and is given by $E(y_1 - y_0|t = 1, X)$, where X is a vector of predictor variables. This could be estimated if the following condition is recognized:

$$E(y_1 - y_0|t = 1, X) = E(y_1|t = 1, X) - E(y_0|t = 1, X) \quad (1)$$

$E(y_1|t = 1, X)$ is the average outcome variable for BO-backed IPOs. However, $E(y_0|t = 1, X)$, the average outcome variable that BO-backed IPOs would experience if they did not receive BO backing, is unobservable. The traditional approach is to use $E(y_0|t = 0, X)$ instead, the average outcome variable of stand-alone IPOs. Unfortunately, because BO backing is not randomly assigned but represents an endogenous choice, this creates a bias. The bias is defined as:

$$B(X) = E(y_0|t = 1, X) - E(y_0|t = 0, X) \quad (2)$$

Rubin (1974, 1977) and Rosenhaum and Rubin (1983) show that, under certain conditions, matching on $\Pr(t = 1|\, X)$ eliminates the bias and reduces the dimensionality of the problem. I use two widely used matching estimator techniques to do that: (i) one-to-one nearest neighbor matching (Rubin, 1974, 1977; Rosenbaum and Rubin, 1983) and (ii) one-to-many smoothed weighted matching (Heckman et al., 1997).

With the one-to-one nearest neighbor matching techniques the distance between pairs of observations with regard to a set of covariates is calculated. Then each observation of the treatment group is matched to a comparable observation of the control group that is closest to it. This technique has two different types: propensity score matching; and the full covariate matching. Propensity score matching requires the estimation of a model for the endogenous choice treatment variable with a vector of X observable variables. The predicted probability of treatment is then used as the propensity score, and each observation of the treatment group is matched with an observation of the control group with the highest propensity score. The full covariate matching does not require the estimation of a formal treatment model. The nearest neighbor is instead determined by using a weighted function of the covariates for each observation.

The one-to-many smoothed weighted matching techniques use a weighted average of the outcomes of several (or perhaps all) observations of the control group to calculate the treatment effect. The weight given to each observation of the control group is in proportion to "closeness" of the vector of X observable variables. Again, there are two main types of this technique, the regression-adjusted local linear matching and the inverse probability weighing matching. The regression-adjusted (RA) local linear matching performs a linear regression to predict potential outcomes from observable variables. To estimate the treatment effect, local weights are calculated, with more weight given to outcomes of observations of the control group that are similar in the predicted outcomes to those of the treatment group. The inverse probability weighting (IPW) matching performs a regression to predict the probability of treatment, instead of potential outcomes, from observable variables. To estimate the treatment effect, local weights are calculated, with more weight to outcomes of observations of the control group with high probability of treatment. So the RA builds a formal model for the outcome, whereas the IPW builds a formal model for the treatment status.

I do all matching with replacement and rely on Abadie and Imbens standard errors to conduct statistical inference because they are more appropriate than bootstrapped standard errors for matching estimators (Abadie and Imbens, 2006). Finally, I calculate selection-bias-adjusted 95% confidence intervals.

4.4 The Results

4.4.1 Pre-IPO Default Risk, Fundamentals and IPO Timing

In this section, I examine the pre-IPO period in order to understand and compare IPO timing considerations for both PE-backed and stand-alone IPOs. Initially, I apply the above matching approach and test whether PE-backed IPOs have higher or lower default risk in the year prior to the IPO compared to the default risk of comparable stand-alone IPOs. Then, I move to a regression framework and test whether PE sponsors or insiders of stand-alone companies are more likely to time their IPOs for when overall market conditions are more favorable.

4.4.1.1 Do PE sponsors time their IPOs when their default risk is lower compared to matched companies?

Here, I first examine various fundamentals in the year prior to the IPO that are related to default risk (such as profitability and leverage) then calculate default measures and, finally, perform comparisons between PE-backed and stand-alone companies. Tables 4.2 and 4.3 show average selection-bias-adjusted default measures and financial ratio differences between BO-backed versus stand-alone IPOs and VC-backed versus stand-alone IPOs respectively. I use two one-to-one nearest neighbor techniques (propensity score and full covariate matching) and two one-to-many smoothed weighted techniques (regression-adjusted local linear and inverse probability weighting matching). Abadie and Imbens (2006) standard errors appear in parentheses and 95% confidence intervals are in square brackets.

The default measures that I use are Altman and Hotchkiss's (2005) survival probability (Z''score), Zmijewski's (1984) default probability, and Shumway's (2001) default probability, calculated from Chava and Jarrow's (2004) model for private firms (who re-estimate Shumway's (2001) model augmented with industry and interaction terms). The financial ratios that I choose to present here as fundamentals are the leverage ratio (total assets to total liabilities), the earnings per share (net income to shares outstanding) and the return on assets (net income to total assets).

In both Tables 4.2 and 4.3, Panel A presents my basic model, in which I use the following dummies as instruments in the first-stage regression that predicts either BO backing (Table 4.2) or VC backing (Table 4.3): the underwriter rank, the logarithm of total assets in the year prior to the IPO, two-digit SIC dummies, headquarter-state dummies, offer year dummies, the natural logarithm of firm age, the number of total managers, road show success dummies, and stock exchange . Jain and Kini (2000) show that PE-backed IPOs attract more prestigious underwriters, have a higher number of total managers and greater road show success. Thus, it is important to use these variables as instruments in my first-stage regression. To measure road show success, I construct three road show success dummies based on the offer price being below, within, or above the initial filing range. Moreover, I use total assets and age as instruments, since the summary statistics in Table 4.1 show substantial differences in these variables. Also evident from Table 4.1 are the industry and geographic concentrations of PE-backed IPOs. Thus SIC and headquarter-state dummies are necessary instruments. Time-series variation in IPOs and the presence of "hot" markets suggest that I should control for year effects as well. Finally, I use stock exchange dummies because there are differences in listing restrictions among stock exchanges. Although I use the same instruments to predict BO and VC backing, I estimate two different models because the two kinds of PE sponsor invest in different companies.

Table 4.2 reports results for BO-backed IPOs. In Panel A, I find that BO-backed IPOs have significantly lower survival probabilities in the year prior to the IPO, as measured by the Altman and Hotchkiss (2005) model, and higher default probabilities (as measured by Zmijewski's (1984) model) compared to matched stand-alone IPOs. Specifically, BO-backed IPOs have significantly lower survival probabilities (3.2% to 5%) and significantly higher default probabilities (4.8% to 8.5%). Shumway's (2001) model gives insignificant results. I also find that they have significantly higher leverage ratios (12% to 15%), but the evidence for differences in other fundamentals is rather weak. Thus, in Panel B, I add the leverage ratio as an instrument and, interestingly, differences based on Altman and Hotchiss (2005) and Zmijewski's (1984) models are no longer significant, whereas according to Shumway's (2001) model, BO-backed IPOs have significantly lower default probabilities pre-IPO this time 2.4% to 4.2%. This is evidence that the differences in the default measures of Panel A are mostly driven by the higher leverage ratios of BO-backed IPOs and do not hold when I compare firms with similar leverage levels. In Panels C and D, I repeat the analysis performed in Panels A and B, controlling for additional instruments. Specifically, I include an all-star analyst dummy, a syndicate dummy and the overallotment percentage. Again, moving from Panel C to Panel D, I add the leverage ratio as an instrument and I no longer find evidence that the pre-IPO default measures of BO-backed and matched stand-alone IPOs differ.

Table 4.3 reports results for VC-backed IPOs. Similarly to the results for BO-backed IPOs, in Panel A, I find that VC-backed IPOs have significantly lower survival probabilities pre-IPO (as measured by the Altman and Hotchkiss (2005) model) and higher default probabilities (as measured by both Zmijewski's (1984) and Shumway's (2001) models) compared to matched stand-alone IPOs. Specifically, VC-backed IPOs have significantly lower survival probabilities (5.2% to 14%) and significantly higher default probabilities (1.8% to 8.3%). Contrary to the findings on BO-backed IPOs, though, these differences are not due to higher leverage ratios of VC-backed IPOs compared to matched stand-alone IPOs, but rather due to significant differences in their profitability ratios. Specifically, VC-backed IPOs have lower earnings per share and ROA (14.5% to 22.7%) than matched stand-alone IPOs. Thus, in Panel B, I add the ROA as an instrument and, interestingly, almost all differences lose their significance. This is evidence that the differences in the default measures of Panel A are mostly driven by the lower profitability ratios of VC-backed IPOs and do not hold when I compare firms with similar profitability. As before, in Panels C and D I repeat the analysis performed in Panels A and B, controlling for additional instruments and find the same results.

In unreported results, I find that results for both BO and VC-backed IPOs are robust to different default measures such as Olson's O-score (1980), different measures of size (i.e. number of employees, total sales or IPO proceeds instead of total assets) and different ratios.

To sum up, as expected, I find that in the year prior to the IPO, BO-backed companies have significantly higher leverage ratios and VC-backed companies have significantly lower profitability ratios than matched stand-alone companies. These differences explain entirely the greater prevalence of measures indicating default for both BO and VC-backed IPOs in the year prior to the IPO. Thus, I do not find strong evidence of a difference between pre-IPO default measures of PE-backed and matched stand-alone IPOs differ after I control for leverage, in the case of BO-backing, and profitability, in the case of VC-backing. I interpret this as evidence that PE sponsors do not time their IPOs in years that they have lower default risk, thus better financial situation, compared to their peers.

Even if PE sponsors are not more proficient than stand-alone companies in timing their IPOs for when company's financial situation is better compared to peers, it can be that they are better in timing the IPO when overall market conditions are more favorable. This is another version of IPO timing that I test below.

4.4.1.2 Are PE sponsors better than insiders of stand-alone companies in timing their IPOs when overall market conditions are more favorable?

To answer this question, I move to a regression framework and look at both the exact period as well as the age of the firm when IPOs take place. Firstly, I study whether PE-backed IPOs are more common than stand-alone IPOs in hot market periods. Secondly, I study if PE-backed IPOs that take place in hot market periods are younger firms at the time of the IPO compared to stand-alone IPOs.

Table 4.4 reports the regression results of two IPO timing measures on a BO and a VC dummy, along with controls that include firm and IPO specific characteristics. I construct the two dependent variables as follows. The first (columns 1-3) is a dummy equal to one if the average first-day return in the month of the IPO is above the median for the period January 1975 to August 2014 (12.5%). The second (columns 4-6) is a dummy equal to one if the percentage of

IPOs with an offer price above the midpoint of the initial offer file range in the month of the IPO is above the median for the period January 1980 to August 2014 (42%). The coefficients of interest are the ones of the BO and VC dummies. The coefficient of the BO dummy is negative and significant in four out of the six specifications. This is some evidence that BO-backed IPOs are less common than other IPOs in hot market periods. The coefficient of the VC dummy is positive and significant in only one specification. Given the above results, I cannot argue that PE sponsors target their IPOs in hot market periods more than stand-alone companies do. In the case of BO sponsors, there is indeed some evidence to the contrary. Other results that demonstrate some significance are that IPOs that have negative EPS in the year prior the IPO and IPOs that receive coverage from an all-star analyst are more likely to take place during hot market periods.

Table 4.5 reports the regression results of firm age at the time of the IPO on a BO and a VC dummy, two IPO market hotness measures, and interactions among these measures and the BO and VC dummies, along with controls that include firm and IPO specific characteristics. To proxy for market conditions at the time of the IPO, I follow Cao (2011) and construct two measures. The first (column 1) is the average market first day return in the three months prior to the IPO and the second (column 2) is the number of IPOs in the three months prior to the IPO. As expected, I find that BO-backed IPOs tend to occur for older firms and VC-backed IPOs for younger firm on average. This is consistent with the summary statistics in Table 4.1 (Panel C). In specification (2), I also find some evidence that companies enter at a younger age when there is a high number of IPOs in previous months. This is in accordance with the intuition that companies rush into IPOs when the market conditions are favorable, and thus they can take advantage of higher valuations. The coefficients of the interactions terms of BO and VC dummies, which are of interest, are negative and significant though (at 10% and 5% respectively). This indicates that the firms that undergo BO and VC-backed IPOs during hot market periods are older on average than stand-alone IPOs that take place during the same periods. Thus, I can argue that, if anything, PE sponsors are less likely to rush their companies into premature IPOs. Other results that demonstrate significance are that companies which are older at the time of the IPO are also more levered, more profitable and larger.

4.4.2 IPO Underpricing and Valuations

In this section, I test whether PE sponsors, as professional insiders, take more advantage of information asymmetries during the IPO compared to stand-alone companies. To test this, I first follow the literature and measure differences in information asymmetries through differences in underpricing. Then, I look directly at the proceeds from IPOs and examine whether PE sponsors inflate valuations more or less compared to similar stand-alone companies.

Tables 4.6 and 4.7 show the average differences (both for selection-bias-adjusted first-day returns and the proceeds from IPO) between VC-backed versus stand-alone IPOs, and BO-backed versus stand-alone IPOs respectively using the methods described above. As before, in my basic model, I use the following instruments in the first-stage regression that predicts either VC backing (Table 4.6) or BO backing (Table 4.7): the underwriter rank, the logarithm of total assets, two-digit SIC dummies, headquarter-state dummies, offer year dummies, the natural logarithm of firm age, the number of total managers, road show success dummies, and stock exchange dummies. . In Panels B, C, and D, I include additional instruments to control for differences in the equity, default probability and leverage ratio, also obvious from Panel C of Table 4.1. Given the missing value problem associated with some of these data items and the different methods employed, the observations used in the estimation are not always equal among panels and methods.

Table 4.6 reports results for VC-backed IPOs. In accordance with the literature (Gompers, 1996; Lee and Wahal, 2004; Neus and Walz, 2005) and in line with the summary statistics in Table 4.1 (Panel C), I find some evidence that first-day returns of VC-backed IPOs are higher compared to stand-alone IPOs. These studies suggest that VC-backed companies face higher information asymmetries in the IPO market. Thus, VC sponsors are willing to underprice IPOs to signal their quality and use underpricing as a mechanism to establish their reputation. Specifically, I find that first-day returns of VC-backed IPOs are significantly higher by 4.7% to 7.2% (Panel D, inverse probability weighting and regression-adjusted local linear matching) when I use the one-to-many smoothed weighted techniques. As a next step, I go beyond existing literature and look directly at the size of the IPO. In the vast majority of cases, I do not observe significant differences in net IPO proceeds between VC-backed and matched stand-alone IPOs. Combining the above results, VC-backed companies have similar initial valuations with stand-alone firms but their stocks often have higher first-day returns. This is an indication that the market soon recognizes them to be “stars”.

Now I test if these results differ in “hot” versus “cold” IPO valuation periods. For this reason, I split my sample in two sub-periods based on Jay Ritter’s monthly measure of "hotness" of the IPO market: the percentage of deals that are priced above the midpoint of the initial file price-range. Specifically, I characterize high valuation periods as those months when the percentage of IPOs with an offer price above the midpoint of the initial offer range is above the median for January 1980 to August 2014 (42%). I characterize low valuation periods as months when this percentage is below median. The two sub-periods are almost equal in size (425 “high valuation" months and 418 "low valuation" months). For the sub-period analysis, I do not calculate sub-period estimates by partitioning the full sample results because they are obtained from a first-stage regression that uses all the data. Instead, I follow Lee and Wahal (2004) and re-estimate the first-stage regression for each sub-period, thereby I tighten the conditioning information and find more conservative estimates.

As before, I find the differences in first-day returns to be almost always positive and both economically and statistically significant in many cases, especially during low valuation periods. In such periods, the differences demonstrate significance under all techniques and are larger in size, ranging from 3.7% (Panel D, full covariate matching) to 12.4% (Panel C, propensity score matching). This is evidence that when valuations are low, VC-backed IPOs outperform stand-alone IPOs even more, as they have quite high first-day returns. Differences of net proceeds remain insignificant in both "hot" and "cold" markets.

Table 4.7 reports results for BO-backed IPOs. First-day returns in most cases do not differ significantly between BO-backed and stand-alone IPOs. This indicates that information asymmetry is similar for both. Interestingly, I always find that BO-backed IPOs have lower valuations as measured by the net IPO proceeds, compared to matched stand-alone IPOs. Results are both economically and statistically significant, as well as robust to the use of the book value of equity, default probability and float percentage as instruments (Panels A, C and D respectively), which suggests that the lower valuations of BO-backed IPOs are not solely driven by their higher indebtedness/default risk levels and potentially lower float percentages. This is an unexpected finding, since, in the summary statistics in Table 4.1 (Panel C), BO-backed IPOs have more than double the net proceeds of stand-alone IPOs. Specifically, I find that IPO proceeds for BO-backed IPOs are significantly lower for the full sample by 29.51 (Panel B, full covariate matching) to 205.71 million (Panel A, inverse probability weighting matching). This is evidence that BO sponsors do not exploit information asymmetries in the IPO market to inflate valuations of their portfolio companies. On the contrary, public investors value these IPOs more conservatively. All findings hold in both high and low valuation periods.

In unreported results, I find that results for both VC and BO-backed IPOs are robust to different measures of size, to gross instead of net proceeds, to addition of other variables such as the initial amount filed, a positive earnings dummy, an "all-star" analyst dummy and a syndicate dummy as instruments, and to another measure of IPO "hotness" (based on the average first day return being above the median for the period January 1975 to August 2014, which is 12.5%).

To sum up, VC-backed IPOs have higher first-day returns than matched stand-alone IPOs, a result that is in accordance with the literature and indicates that VC sponsors are confronted with higher information asymmetries in the IPO market. BO-sponsors on the contrary have similar underpricing as their peers, i.e. information asymmetries between BO sponsors and public investors are similar to the asymmetries between managers and public investors. Finally, neither VC nor BO-backed companies have inflated valuations compared to similar companies. In the case of BO sponsors, there is indeed evidence that they are conservatively priced. The reason for this may be that investors perceive them as riskier due to their high leverage. In the next section, I test if this fear is justified, specifically if they default more often post-IPO.

4.4.3 Post-IPO Default and Failure Risk

In the sections above, I concentrate on the pre-IPO and IPO period. It may be that PE investors strategically choose to exit via an IPO before hidden problems in their portfolio companies unfold. Thus, what is of utmost interest is the question of whether BO and VC-backed companies delist more often than stand-alone companies for reasons related to default or failure. I am interested in two types of delistings: delistings due to default and, more generally, delistings due to failure. I define as default cases of bankruptcy or liquidation (codes 400-490 and 574). I define as failure cases of bankruptcy, liquidation, or delisting due to other negative reasons such as non-payment of fees (codes 400-490, 550-561 and 574-591).

Tables 4.8 and 4.9 show average selection-bias-adjusted default and failure rates differences between BO-backed versus stand-alone IPOs and VC-backed versus stand-alone IPOs respectively using the methods described above. Here, I track companies for five years after the IPO, thus I do not examine IPOs after 2009. By right censoring my data I control for survivorship bias by which older firms undergoing IPOs provide more years of data and have a longer period "at risk". I examine a horizon of five years for all companies to correct for this effect. Robustness tests with different horizons provide qualitatively the same results. A shorter horizon though, (e.g. for one year) may be inappropriate since IPOs generally feature lockup provisions that prohibit corporate insiders from selling shares before a certain date. In our sample lockup provisions for PE-backed IPOs are no different from those of other IPOs and have an average of six months. These lockup provisions help align the interests of insiders with those of public investors (Field and Hanka, 2001, Aggarwal et al., 2002) and may decrease default and failure risk during this period. As before, in my basic model, I use the following as instruments in the first-stage regression that predicts either BO backing (Table 4.8) or VC backing (Table 4.9): the underwriter rank, the logarithm of total assets, two-digit SIC dummies, headquarter-state dummies, offer year dummies, the natural logarithm of firm age, the number of total managers, road show success dummies and stock exchange dummies.

Table 4.8 reports results for BO-backed IPOs. In Panel A, I find that default and failure rates of BO-backed and stand-alone IPOs do not differ significantly. The same holds in Panel B, where I add the leverage ratio as an instrument. In Panels C and D I repeat the analysis performed in Panels A and B, controlling for additional instruments. Specifically, I include an all-star analyst dummy, a syndicate dummy and the overallotment percentage. Again, moving from Panel C to Panel D, I add the leverage ratio as an instrument and find the same results.

Table 4.9 reports results for VC-backed IPOs. As for BO-backed IPOs, I find that, in most cases, default and failure rates of VC-backed and stand-alone IPOs do not differ significantly. There is some weak evidence that VC-backed IPOs have lower failure rates than stand-alone IPOs since differences are negative and significant in a few cases, ranging from 2.8% (Panel A, inverse probability weighting matching) to 5.3% (Panel B, inverse probability weighting matching).

In order to shed more light on the above results and further test if PE-backed IPOs delist more often than stand-alone IPOs due to default or failure, I now move to a regression framework where I use all firm-year observations post-IPO. Specifically, I apply the multi-period logit regression framework (Shumway, 2001) for my default and failure prediction models. As described above, I use two dependent variables in my regressions. The first (Table 4.10) is a default dummy equal to one in the year that the company delists due to bankruptcy or liquidation (codes 400-490 and 574). The second (Table 4.11) is a failure dummy equal to one in the year that the company delists due to bankruptcy, liquidation, or due to other negative reasons (codes 400-490, 550-561 and 574-591). I regress these dummies on a BO and a VC dummy, along with various firm and IPO specific characteristics. To proxy for market conditions at the time of the IPO, I use the same measures as in Table 4.7 and examine interaction of these variables with the

BO and VC dummies to additionally test if BO and VC-backed IPOs that take place in hot market periods are more likely to default or fail afterwards.

Table 4.10 reports the regression results for the default dependent variable from four different specifications. My coefficients of interest are those of the BO and VC dummies which are negative and significant in three out of the four specifications. This is evidence that BO and VC-backed IPOs default less than others. IPOs that take place in hot market periods (i.e. in months that follow high IPO activity as measured by the number of IPOs taking place in the previous three months) are significantly more likely to default on average. This is consistent with my previous finding in Table 4.5 that companies often rush into premature IPOs when market conditions are favorable. It is also in accordance with the theoretical model of Yung, Colak and Wang (2008), according to which there are more delistings in hot market periods. Since the coefficients of the interaction terms in column (3) are not significant though, I cannot argue that this result is stronger for either BO or VC-backed IPOs. As expected, companies with lower profitability and higher leverage are more likely to default as indicated from the signs of the profitability and leverage ratios. Table 4.11 reports the regression results for the failure dependent variable using the same four specifications. When we move from Table 4.10 to Table 4.11, results remain substantially similar. The above results indicate that PE sponsors are not any more likely than managers of stand-alone companies to "unload lemons" in the IPO market and it is in accordance with the intuition that if these sponsors are caught "cheating", they will struggle to raise money for future funds.

4.5 Concluding Remarks

I study the role of both BO and VC sponsors in a setting of high information asymmetry as the IPO market. These professional insiders may be more capable of taking advantage of such asymmetries compared to insiders of stand-alone companies. But they also have more reputational capital at stake, a factor which tends to be known by the market. BO and VC sponsors may also behave differently from each other. Thus, I differentiate my analysis for each type of PE sponsor and compare BO and VC-backed IPOs with IPOs of stand-alone companies in a matching framework.

I do not find significant differences between these IPOs and matched IPOs of stand-alone companies. The financial situation of both BO and VC-backed companies in the pre-IPO year, as measured by their default risk, is similar to that of their peers. Moreover, PE sponsors do not target their IPOs in hot periods any more than do managers of stand-alone companies. They also are not more prone to rush their companies into premature IPOs and do not inflate valuations. Finally, PE-backed companies do not default more often post-IPO. This is evidence that PE sponsors are not more likely to seek to sell firms with poor prospects ("unload lemons") in the IPO market.

This paper provides evidence against the criticism that PE sponsors often receive in the media (e.g. "Rush to get to the front of the IPO queue" in *Financial Times*, 18 February 2014). It can also have important policy implications on the regulatory framework related to PE, such as the Dodd-Frank Act (signed by Obama in July 2010). Finally, it comes as a timely contribution, given the increasing importance of PE-backed IPOs in the market ("Private equity-backed IPOs could hit seven-year high" in the *Financial Times*, 29 September 2014).

4.6 Tables of Chapter 4

Table 4.1

Distribution and Characteristics of IPOs, 1975 to 2014

The table shows the distribution and characteristics of 7,033 IPOs for which data are available after excluding those with an offer price below $5.00 per share, a size below 1.5 million, unit offers, American Depository Receipts (ADRs), closed-end funds, natural resource partnerships, acquisition companies, Real Estate Investment Trusts (REITs), bank and Savings and Loans (S&L) IPOs. I classify growth-capital backed IPOs as VC-backed. Panel A shows the distribution of all IPOs, BO-backed IPOs and VC-backed IPOs across two-digit SIC Codes both as a number and a percentage of all IPOs in each SIC Code. SIC Codes in which there are less than 30 IPOs over the entire sample period are not shown. Panel B shows the geographic distribution of all IPOs, BO-backed IPOs and VC-backed IPOs both as a number and a percentage of all IPOs headquartered in each state. States with less than 140 IPOs, corresponding to 2% of all IPOs, are not shown. Finally, Panel C provides means of various characteristics of stand-alone IPOs, BO-backed IPOs and VC-backed IPOs. Gross and net proceeds are in millions of dollars. The gross spread is in percent. Underwriter rankings are on a 0 to 9 scale, with higher ranking to more prestigious underwriters. All financial statements data are from up to two years prior to the offering. Assets, sales and book and market value of equity are in millions of dollars. The market value is calculated using the post issue shares outstanding multiplied by the offer price. I calculate first-day returns as the percentage price movement from the offer price to the close price on the first trading day. Age is the average number of years from the founding date to the IPO date. The default rate is the percentage of delistings due to bankruptcy and liquidation (CRSP delisting codes 400-490 and 574).

	All IPOs	BO-backed		VC-backed			All IPOs	BO-backed		VC-backed	
	#	#	%	#	%		#	#	%	#	%
Panel A. Distribution by two-digit SIC Code											
15	39	6	15.38	1	2.56	49	81	6	7.41	19	23.46
20	108	24	22.22	14	12.96	50	186	26	13.98	44	23.66
23	68	12	17.65	8	11.76	51	82	13	15.85	11	13.41
25	31	8	25.81	4	12.90	53	33	13	39.39	6	18.18
26	37	13	35.14	7	18.92	54	51	23	45.10	4	7.84
27	64	16	25.00	10	15.63	56	81	21	25.93	15	18.52
28	553	45	8.14	356	64.38	57	71	15	21.13	13	18.31
30	57	11	19.30	5	8.77	58	156	29	18.59	33	21.15
33	79	19	24.05	11	13.92	59	172	25	14.53	60	34.88
34	66	21	31.82	7	10.61	63	202	33	16.34	19	9.41
35	453	39	8.61	217	47.90	65	40	5	12.50	4	10.00
36	627	75	11.96	321	51.20	70	62	6	9.68	3	4.84
37	117	41	35.04	12	10.26	73	1,499	100	6.67	858	57.24
38	479	31	6.47	269	56.16	78	69	8	11.59	14	20.29
39	74	11	14.86	9	12.16	79	52	7	13.46	1	1.92
42	62	6	9.68	6	9.68	80	244	28	11.48	108	44.26

44	54	9	16.67	2	3.70	82	36	8	22.22	12	33.33
45	55	2	3.64	12	21.82	87	207	24	11.59	78	37.68
47	35	5	14.29	7	20.00						
48	311	38	12.22	131	42.12	**Full sample**	**7,033**	**897**	**12.75**	**2,763**	**39.29**
	All IPOs	BO-backed		VC-backed			All IPOs	BO-backed		VC-backed	
	#	#	%	#	%		#	#	%	#	%
Panel B. Distribution by State											
California	1,694	130	7.67	1,013	59.80	New Jersey	281	36	12.81	93	33.10
Colorado	147	12	8.16	67	45.58	New York	522	77	14.75	126	24.14
Connecticut	152	25	16.45	53	34.87	Ohio	149	38	25.50	26	17.45
Florida	306	35	11.44	68	22.22	Pennsylvania	252	36	14.29	93	36.90
Georgia	182	28	15.38	75	41.21	Texas	491	62	12.63	163	33.20
Illinois	253	62	24.51	67	26.48	Virginia	162	33	20.37	50	30.86
Massachusetts	481	31	6.44	306	63.62	Washington	149	9	6.04	86	57.72
Minnesota	148	24	16.22	62	41.89	**Full sample**	**7,033**	**897**	**12.75**	**2,763**	**39.29**

Panel C. Characteristics

	Stand-alone IPOs		BO-backed		VC-backed	
	#	Mean	#	Mean	#	Mean
Gross proceeds	3,373	70.63	897	156.85	2,763	58.99
Net proceeds	3,373	65.80	897	145.57	2,763	54.29
Gross spread	3,362	4.30	896	10.15	2,759	4.04
Underwriter rank	3,314	6.38	883	8.30	2,726	7.71
Assets	2,945	725.23	861	836.03	2,563	60.71
Sales	2,460	300.70	773	742.03	2,319	52.17
% profitable	2,986	72.37	846	66.78	2,653	49.19
Book value of equity	3,218	96.50	881	79.44	2,698	21.94
Market value of equity	2,762	338.55	825	661.06	2,533	406.53
Debt/Assets	2,926	0.80	859	0.87	2,559	0.69
Return on assets	2,459	-0.10	773	-0.01	2,317	-0.39
Average first-day return	3,044	14.09	883	10.51	2,588	27.86
Age	3,325	16.19	892	34.20	2,758	8.70
Default rate	3,373	3.91	897	2.34	2,763	2.39

Table 4.2

Pre-IPO selection bias adjusted default measures' and financial ratios' differences between BO-backed and stand-alone IPOs

The table presents selection bias adjusted average default measures' and financial ratios' differences between BO-backed and stand-alone IPOs. Each BO-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of BO-backed IPOs matched appear below the average differences. When BO-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. All financial statements data are from the year prior to the offering. I calculate Altman and Hotchkiss's survival probability from Altman and Hotchkiss's (2005) model. I calculate Zmijewski's default probability from Zmijewski's (1984) model. I calculate Shumway's default probability from Chava and Jarrow's (2004) model for private firms, who re-estimate Shumway's (2001) variables' coefficients augmented with industry and interaction terms. The leverage ratio is the ratio of total assets to total liabilities. The earnings per share is the ratio of net income divided by the number of pre-issue shares outstanding. The Return on Assets is the ratio of net income to total assets. ** denotes significance at a 5% level and * at a 10% level.

Δ(Probabilities)				Δ(Ratios)			
One-to-one nearest neighbor		One-to-many smoothed weighted		One-to-one nearest neighbor		One-to-many smoothed weighted	
Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Altman and Hotchkiss's (2005) survival probability				Leverage ratio			
-0.050	**-0.047**	**-0.032**	**-0.049**	**0.135**	**0.139**	**0.119**	**0.150**
(-2.56)**	**(-3.30)****	**(-2.59)****	**(-4.65)****	**(4.27)****	**(5.54)****	**(5.35)****	**(7.66)****
[-0.088,-0.012]	[-0.075,-0.019]	[-0.057,-0.008]	[-0.069,-0.028]	[0.073,0.198]	[0.090,0.188]	[0.075,0.162]	[0.112,0.189]
698	453	699/2,810	699/2,810	841	613	842/3,691	842/3,691
Zmijewski's (1984) default probability				Earnings per share			
0.048	**0.050**	**0.049**	**0.085**	-2.344	-1.411	-2.884	-3.118
(2.82)**	**(3.24)****	**(4.09)****	**(5.51)****	(-1.42)	(-1.46)	**(-1.83)***	**(-1.92)***
[0.015,0.082]	[0.020,0.080]	[0.025,0.072]	[0.038,0.079]	[-5.579,0.890]	[-3.301,0.479]	[-5.975,0.206]	[-6.294,0.057]
732	516	733/3,034	733/3,034	826	604	827/3,723	827/3,723
Shumway's (2001) default probability				Return on Assets			
-0.005	0.020	-0.009	-0.001	-0.017	-0.008	0.011	-0.028
(-0.56)	(1.36)	(-1.40)	(-0.20)	(-0.62)	(-0.30)	(0.44)	**(-2.20)****
[-0.021,0.012]	[-0.009,0.048]	[-0.022,0.004]	[-0.008,0.006]	[-0.073,0.038]	[-0.058,0.043]	[-0.036,0.027]	[-0.054,-0.003]
754	516	755/3,144	755/3,144	755	538	756/3,150	756/3,150
Panel B. Instrumental variables: Underwriter rank, log (total assets), leverage ratio, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Altman and Hotchkiss's (2005) survival probability							
-0.024	**-0.033**	-0.001	-0.028				
(-1.20)	**(-1.70)***	(-0.08)	(-1.09)				
[-0.062,0.0149]	[-0.070,0.005]	[-0.024,0.022]	[-0.077,0.022]				
698	453	699/2,810	699/2,810				
Zmijewski's (1984) default probability							
0.013	0.004	0.005	**0.016**				
(0.66)	(0.42)	(0.59)	**(1.66)***				

Table 4.2. Panel B. Cont.

[-0.026,0.053]	[-0.016,0.025]	[-0.012,0.022]	[-0.003,0.036]
732	516	733/3,034	733/3,034
	Shumway's (2001) default probability		
-0.024	0.000	**-0.042**	**-0.027**
(-2.14)**	(0.06)	**(-2.14)****	**(-4.29)****
[-0.046,-0.002]	[-0.022,0.023]	[-0.054,-0.033]	[-0.039,-0.015]
754	516	755/3,144	755/3,144

Panel C. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage

	Altman and Hotchkiss's (2005) survival probability				Leverage ratio		
-0.078	**-0.042**	**-0.038**	**-0.052**	**0.154**	**0.131**	**0.096**	**0.151**
(-3.66)**	**(-2.97)****	**(-2.60)****	**(-4.06)****	**(3.31)****	**(5.01)****	**(3.30)****	**(5.50)****
[-0.120,-0.036]	[-0.069,-0.141]	[-0.066,-0.009]	[-0.078,-0.027]	[0.063,0.245]	[0.080,0.182]	[0.039,0.153]	[0.097,0.205]
448	453	450/1,852	450/1,845	512	613	474/2,304	474/2,010
	Zmijewski's (1984) default probability				Earnings per share		
0.059	**0.046**	**0.044**	**0.060**	-2.27	-1.44	-2.72	-2.16
(2.59)**	**(3.01)****	**(2.97)****	**(4.46)****	(-0.92)	(-1.50)	(-1.33)	(-1.36)
[0.014,0.103]	[0.162,0.076]	[0.015,0.073]	[0.034,0.086]	[-7.103,2.557]	[-3.325,0.446]	[-6.730,1.289]	[-5.276,0.959]
472	516	474/2,017	474/2,010	503	604	461/2,253	461/2,010
	Shumway's (2001) default probability				Return on Assets		
-0.013	-0.011	-0.014	-0.003	-0.032	0.014	0.010	-0.021
(-1.00)	(-1.33)	(-1.60)	(-0.49)	(-0.76)	(0.49)	(0.36)	(-1.10)
[-0.039,0.013]	[-0.027,0.005]	[0.031,0.003]	[-0.014,0.008]	[0.0627,0.245]	[-0.043,0.0720]	[-0.046,0.067]	[-0.058,0.016]
479	537	474/2,055	474/2,010	480	538	474/2,060	474/2,010

Panel D. Instrumental variables: Underwriter rank, log (total assets), leverage ratio, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage

	Altman and Hotchkiss's (2005) survival probability		
-0.021	-0.019	-0.007	**-0.024**
(-0.83)	(-1.46)	(-0.52)	**(-1.89)***
[-0.069,0.028]	[-0.044,0.006]	[-0.034,0.020]	[-0.049,0.001]
448	453	450/1,852	450/1,845
	Zmijewski's (1984) default probability		
0.014	0.004	0.006	0.017
(0.49)	(0.33)	(0.59)	(1.43)
[-0.042,0.070]	[-0.018,0.025]	[-0.014,0.026]	[-0.006,0.040]
472	516	474/2,017	474/2,010
	Shumway's (2001) default probability		
-0.016	-0.009	**-0.044**	**-0.031**
(-1.11)	(-1.44)	**(-2.25)****	**(-3.57)****
[-0.044,0.012]	[-0.022,0.003]	[-0.058,-0.030]	[-0.048,-0.014]
479	537	474/2,055	474/2,010

Table 4.3

Pre-IPO selection bias adjusted default measures' and financial ratios' differences between VC-backed and stand-alone IPOs

The table presents selection bias adjusted average default measures' and financial ratios' differences between VC-backed and stand-alone IPOs. Each VC-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of VC-backed IPOs matched appear below the average differences. When VC-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. All financial statements data are from the year prior to the offering. I calculate Altman and Hotchkiss's survival probability from Altman and Hotchkiss's (2005) model. I calculate Zmijewski's default probability from Zmijewski's (1984) model. I calculate Shumway's default probability from Chava and Jarrow's (2004) model for private firms, who re-estimate Shumway's (2001) variables' coefficients augmented with industry and interaction terms. The leverage ratio is the ratio of total assets to total liabilities. The earnings per share is the ratio of net income divided by the number of pre-issue shares outstanding. The Return on Assets is the ratio of net income to total assets. ** denotes significance at a 5% level and * at a 10% level.

Δ(Probabilities)				Δ(Ratios)			
One-to-one nearest neighbor		One-to-many smoothed weighted		One-to-one nearest neighbor		One-to-many smoothed weighted	
Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Altman and Hotchkiss's (2005) survival probability				Leverage ratio			
-0.131	**-0.098**	**-0.052**	**-0.140**	-0.060	**-0.056**	**-0.067**	-0.039
(-5.25)**	**(-4.88)****	**(-2.88)****	**(-4.93)****	(-1.62)	**(-2.08)****	**(-3.20)****	(-1.43)
[-0.181,-0.082]	[-0.137,-0.059]	[-0.088,-0.017]	[-0.196,-0.084]	[-0.133,0.0127]	[-0.109,-0.003]	[-0.108,-0.026]	[-0.092,0.0142]
2,234	1,734	2,208/4,537	2,208/4,319	2,529	2,088	2,235/5,380	2,235/4,536
Zmijewski's (1984) default probability				Earnings per share			
0.083	**0.037**	**0.083**	**0.078**	**-1.106**	**-1.142**	**-2.009**	**-1.475**
(3.44)**	**(5.52)****	**(6.81)****	**(4.34)****	**(-1.64)***	**(-1.95)***	**(-2.25)****	**(-3.41)****
[0.035,0.130]	[0.024,0.050]	[0.059,0.107]	[0.043,0.113]	[-2.424,0.213]	[-2.29,0.008]	[-3.356,-0.263]	[-2.323,-0.627]
2,234	1,860	2,235/4,537	2,235/4,536	2,614	2,163	2,202/5,512	2,202/4,371
Shumway's (2001) default probability				Return on Assets			
0.073	0.020	**0.018**	**0.076**	**-0.145**	**-0.182**	**-0.227**	**-0.186**
(3.10)**	(1.36)	**(2.14)****	**(5.47)****	**(-2.72)****	**(-5.00)****	**(-7.82)****	**(-4.75)****
[0.027,0.120]	[-0.009,0.480]	[0.002,0.034]	[0.048,0.103]	[-0.250,-0.041]	[-0.253,-0.110]	[-0.278,-0.167]	[-0.263,-0.109]
2,234	1,806	2,235/4,681	2,235/4,536	2,290	2,088	2,235/4,686	2,235/4,536
Panel B. Instrumental variables: Underwriter rank, log (total assets), Return on Assets, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Altman and Hotchkiss's (2005) survival probability							
-0.069	0.010	-0.020	-0.024				
(-2.73)**	(0.62)	(-1.00)	(-0.77)				
[-0.118,-0.020]	[-0.021,0.040]	[-0.060,0.020]	[-0.084,0.037]				
2,234	1,734	2,208/4,537	2,208/4,319				
Zmijewski's (1984) default probability							
-0.011	0.011	0.013	**-0.042**				
(-0.42)	(0.98)	(1.46)	**(-1.77)***				

Table 4.3. Panel B. Cont.

[-0.062,0.040]	[-0.011,0.032]	[-0.004,0.030]	[-0.088,0.004]				
2,234	1,860	2,235/4,537	2,235/4,536				
	Shumway's (2001) default probability						
-0.025	0.001	**-0.039**	**-0.053**				
(-0.94)	(0.06)	**(-2.92)****	**(-2.15)****				
[-0.078,0.027]	[-0.022,0.023]	[-0.049,-0.030]	[-0.101,-0.005]				
2,234	1,806	2,235/4,681	2,235/4,536				

Panel C. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage

	Altman and Hotchkiss's (2005) survival probability				Leverage ratio		
-0.131	**-0.100**	**-0.051**	**-0.140**	**-0.077**	**-0.057**	**-0.089**	-0.039
(-4.37)**	**(-4.12)****	**(-2.72)****	**(-4.93)****	**(-1.72)***	**(-2.14)****	**(-3.31)****	(-1.43)
[-0.190,-0.072]	[-0.148,-0.053]	[-0.087,-0.014]	[-0.196,-0.084]	[-0.165,0.011]	[-0.110,-0.005]	[-0.141,-0.036]	[-0.092,0.0142]
1,562	1,734	1,565/4,537	1,565/4,319	1,721	2,088	1,585/3,509	1,585/3,119
	Zmijewski's (1984) default probability				Earnings per share		
0.096	**0.039**	**0.091**	**0.078**	**-3.459**	**-0.887**	**-2.463**	**-1.475**
(3.45)**	**(5.92)****	**(5.88)****	**(4.34)****	**(-2.15)****	**(-1.95)***	**(-2.06)****	**(-3.41)****
[0.042,0.151]	[0.026,0.052]	[0.061,0.121]	[0.043,0.113]	[-6.587,-0.311]	[-1.778,0.004]	[-4.810,-0.117]	[-2.323,-0.627]
1,582	1,860	1,585/3,124	1,585/4,536	1,731	2,163	1,579/3,477	1,579/3,119
	Shumway's (2001) default probability				Return on Assets		
0.082	0.015	0.016	**0.019**	**-0.194**	**-0.185**	**-0.253**	**-0.186**
(2.95)**	(1.05)	(1.45)	**(1.96)****	**(-3.23)****	**(-5.17)****	**(-6.95)****	**(-4.75)****
[0.027,0.136]	[-0.013,0.044]	[-0.006,0.037]	[-0.000,0.039]	[-0.312,-0.076]	[-0.256,-0.115]	[-0.324,-0.181]	[-0.263,-0.109]
1,582	1,806	1,585/3,165	1,585/4,536	1,592	1,860	1,585/3,169	1,585/3,119

Panel D. Instrumental variables: Underwriter rank, log (total assets), Return on Assets, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage

	Altman and Hotchkiss's (2005) survival probability		
-0.071	0.018	**0.033**	-0.042
(-2.48)**	(0.95)	**(-1.70)***	(-1.44)
[-0.127,-0.015]	[-0.019,0.054]	[-0.070,0.005]	[-0.098,0.015]
1,562	1,734	1,565/4,537	1,565/4,319
	Zmijewski's (1984) default probability		
0.025	0.013	0.013	**-0.076**
(0.82)	(1.25)	(1.13)	**(-2.35)****
[-0.035,0.086]	[-0.008,0.034]	[-0.009,0.035]	[-0.139,-0.012]
1,582	1,860	1,585/3,124	1,585/4,536
	Shumway's (2001) default probability		
-0.001	-0.005	**-0.050**	**-0.097**
(-0.03)	(-0.42)	**(-2.75)****	**(-4.01)****
[-0.064,0.061]	[-0.028,0.018]	[-0.062,-0.037]	[-0.145,-0.050]
1,582	1,806	1,585/3,165	1,585/4002C536

Table 4.4

Regressions of IPO timing dummy on firm and IPO characteristics

The table shows OLS regressions of two measures of IPO timing on firm and IPO characteristics. The first dependent variable is a dummy equal to one if the average first-day return in the month of the IPO is above the median. The second dependent variable is a dummy equal to one if the percentage of IPOs with an offer price above the midpoint of the initial offer file range in the month of the IPO is above median. I construct these dummies with data from Jay Ritter's website. The sample consists of all IPOs from Table 4.1 with available data. All financial statements data are from the year prior to the offering. NITA is the ratio of net income to total assets. TLTA is the ratio of total liabilities to total assets. Ln (age) is the natural logarithm of the number of years from the founding date to the IPO date. The positive EPS dummy is equal to one if the earnings per share (ratio of net income divided by the number of pre-issue shares outstanding) is above zero. Underwriter rankings are on a 0 to 9 scale, with higher ranking to more prestigious underwriters. The all-star analyst dummy is equal to one if the offer receives coverage of an all-star analyst. Industry fixed effects are based on two-digit SIC codes. ** denotes significance at a 5% level and * at a 10% level.

	Dummy =1 if average 1st day return > median						Dummy =1 if % of IPOs with offer price above midpoint > median					
	(1)		(2)		(3)		(4)		(5)		(6)	
BO dummy	**-0.142***	**(-1.69)**	**-0.226****	**(-2.20)**	**-0.472****	**(-3.78)**	-0.0128	(-0.15)	-0.166	(-1.64)	**-0.347****	**(-2.82)**
VC dummy	**0.166****	**(2.78)**	0.075	(0.97)	-0.0638	(-0.66)	-0.0156	(-0.26)	-0.0816	(-1.07)	-0.0782	(-0.83)
TLTA			**-0.185****	**(-2.40)**	**-0.247****	**(-2.70)**			-0.0254	(-0.34)	0.0480	(0.55)
NITA			**-0.123***	**(-1.65)**	-0.108	(-1.24)			-0.00611	(-0.09)	0.0567	(0.70)
Ln (age)			**0.0620****	**(2.13)**	0.0317	(0.91)			0.0283	(0.99)	0.000606	(0.02)
Positive EPS dummy			**-0.243****	**(-3.05)**	**-0.222****	**(-2.32)**			**-0.146***	**(-1.88)**	-0.0928	(-1.00)
Underwriter rank			**-0.0523****	**(-2.76)**	-0.0205	(-0.85)			-0.0205	(-1.10)	-0.0277	(-1.16)
All-star analyst dummy					**0.284****	**(2.51)**					**0.188***	**(1.77)**
Constant	-0.802	(-1.04)	0.645	(0.85)	1.561	(0.96)	**-1.943****	**(-2.16)**	0.135	(0.18)	1.421	(0.88)
Other controls		Yes		Yes		Yes		Yes		Yes		Yes
Industry FE		Yes		Yes		Yes		Yes		Yes		Yes
Observations		7,033		5,045		3,507		7,033		5,045		3,507

Table 4.5

Regressions of age at IPO on firm and IPO characteristics

The table shows OLS regressions of the age at IPO on firm and IPO characteristics. I calculate age at IPO as the natural logarithm of the number of years from the founding date to the IPO date. The sample includes all IPOs from Table 4.1 with available data. All financial statements data are from the year prior to the offering. NITA is the ratio of net income to total assets. TLTA is the ratio of total liabilities to total assets. To proxy for market conditions at the time of the IPO, I use two measures: The average market first-day return and the logarithm of the number of IPOs in the three months prior the IPO. I construct these measures with data from Jay Ritter's website. I examine interaction effects for the additional impact of BO and VC-backed IPOs. Industry fixed effects are based on two-digit SIC codes. ** denotes significance at a 5% level and * at a 10% level.

	Age at IPO			
	(1)		(2)	
BO dummy	**0.335****	**(4.88)**	-0.389	(-1.10)
VC dummy	**-0.107****	**(-2.18)**	**-0.796****	**(-2.89)**
TLTA	**0.222****	**(7.02)**	**0.223****	**(7.07)**
NITA	**0.165****	**(5.57)**	**0.168****	**(5.68)**
Log (total assets)	**0.223****	**(2.84)**	**0.223****	**(2.77)**
Previous market first-day return	0.110	(0.65)	-0.0306	(-0.22)
Previous market first-day return * BO dummy	-0.280	(-1.11)		
Previous market first-day return * VC dummy	-0.189	(-1.29)		
Log (number of previous IPOs)	-0.0296	(-0.54)	**-0.117***	**(-1.84)**
Log (number of previous IPOs) * BO dummy			**0.141***	**(1.87)**
Log (number of previous IPOs) * VC dummy			**0.134****	**(2.36)**
Constant	0.190	(0.13)	0.383	(0.26)
Other controls		Yes		Yes
Industry FE		Yes		Yes
Observations		5,362		5,362

Table 4.6

Selection bias adjusted first-day returns' and net IPO proceeds' differences between VC-backed and stand-alone IPOs

The table presents selection bias adjusted average net IPO proceeds' and first-day returns' differences between VC-backed and stand-alone IPOs. Each VC-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of VC-backed IPOs matched appear below the average differences. When VC-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. High valuations periods are months when the percentage of IPOs with an offer price above the midpoint of the initial offer range is above median. Low valuations periods are months when this percentage is below median. I take this measure from Jay Ritter's website. Δ(Proceeds) are in millions of dollars and Δ(Returns) in percentage points. ** denotes significance at a 5% level and * at a 10% level.

Δ(Returns)				Δ(Proceeds)			
One-to-one nearest neighbor		One-to-many smoothed weighted		One-to-one nearest neighbor		One-to-many smoothed weighted	
Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Full sample							
0.012	0.023	**0.071**	**0.051**	-11.34	3.52	-5.73	-6.33
(0.36)	(1.01)	**(4.29)****	**(2.32)****	(-0.67)	(0.28)	(-0.61)	(-0.80)
[-0.054,0.078]	[-0.021,0.067]	[0.039,0.104]	[0.008,0.095]	[-44.51,21.82]	[-21.37,28.40]	[-24.20,12.74]	[-21.76,9.10]
2,475	2,034	2,475/5,208	2,474/5,207	2,541	2,081	2,541/5,375	2,541/5,374
High valuations periods							
0.023	0.000	**0.080**	**0.040**	10.79	-5.45	4.76	6.54
(0.44)	(-0.01)	**(2.98)****	(1.12)	(0.43)	(-1.10)	(0.33)	(0.58)
[-0.774,0.123]	[-0.071,0.071]	[0.028,0.133]	[-0.030,0.110]	[-38.01,59.59]	[-15.14,4.23)	[-23.79,33.30]	[-15.54,28.62]
1,425	1,187	1,425/2,897	1,425/2,889	1,470	1,223	1,470/3,000	1,470/2,994
Low valuations periods							
0.063	**0.038**	**0.045**	**0.054**	-13.92	-2.22	-13.94	**-13.36**
(2.37)**	**(1.96)****	**(3.17)****	**(3.13)****	(-1.06)	(-0.42)	(-1.18)	**(-1.99)****
[0.011,0.116]	[-82e-07,0.075]	[0.017,0.074]	[0.020,0.087]	[-39.78,11.93]	[-12.52,8.08]	[-37.08,9.20]	[-26.51,-0.210]
1,027	709	1,043/2,311	1,043/2,311	1,048	716	1,064/2,375	1,064/2,373
Panel B. Instrumental variables: Underwriter rank, log (total assets), log (equity), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Full sample							
0.031	0.036	**0.071**	**0.052**	-10.65	0.67	-5.58	-5.99
(0.94)	(1.52)	**(4.30)****	**(2.40)****	(-0.62)	(0.05)	(-0.59)	(-0.77)
[-0.034,0.097]	[-0.010,0.083]	[0.189,0.243]	[0.009,0.095]	[-44.18,22.87]	[-24.82,26.17]	[-24.08,12.93]	[-21.27,9.28]
2,474	2,033	2,474/5,200	2,474/5,199	2,540	2,080	2,540/5,367	2,540/5,366
High valuations periods							
-0.005	0.007	**0.080**	0.042	6.85	-7.52	5.28	6.65
(-0.10)	(0.19)	**(3.00)****	(1.19)	(0.27)	(-1.52)	(0.36)	(0.59)
[-0.108,0.097]	[-0.064,0.078]	[0.028,0.132]	[-0.029,0.110]	[-42.44,56.14]	[-17.24,2.20]	[-23.40,33.95]	[-15.46,28.75]
1,425	1,187	1,425/2,893	1,425/2,885	1,470	1,223	1,470/2,996	1,470/2,990

Table 4.6. Panel B. Cont.

			Low valuations periods				
0.073	**0.043**	**0.044**	**0.054**	-20.67	-4.16	-14.16	**-13.32**
(2.50)**	**(2.30)****	**(3.08)****	**(3.13)****	(-1.32)	(-0.77)	(-1.21)	**(-1.99)****
[0.016,0.130]	[0.006,0.080]	[0.016,0.072]	[0.020,0.087]	[-51.44,10.10]	[-14.81,6.49]	[-37.05,8.73]	[-26.44,0.20]
1,026	709	1,063/2,307	1,042/2,307	1,047	715	1,063/2,371	1,042/2,369

Panel C. Instrumental variables: Underwriter rank, log (total assets), default probability, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies

			Full sample				
0.010	0.018	**0.069**	0.039	-7.37	3.15	1.25	-3.93
(0.26)	(0.68)	**(3.48)****	(1.16)	(-0.38)	(0.20)	(0.11)	(-0.44)
[-0.067,0.877]	[-0.034,0.071]	[0.030,0.107]	[-0.023,0.105]	[-45.34,30.59]	[-27.63,33.95]	[-21.13,23.62]	[-21.46,13.59]
2,126	1,679	2,126/4,199	2,126/4,199	2,160	1,707	2,160/4,283	2,160/4,283
			High valuations periods				
-0.060	-0.750	**0.071**	0.005	10.83	-6.74	12.07	9.36
(-0.85)	(-0.11)	**(2.27)****	(0.08)	(0.37)	(-1.19)	(0.74)	(0.73)
[-0.198,0.078]	[-13.85,12.34]	[0.010,0.133]	[-0.12,0.13]	[-46.99,68.65]	[-17.87,4.39]	[-20.02,44.17]	[-15.62,34.34]
1,268	524	1,268/2,440	1,268/2,434	1,296	1,044	1,296/2,501	1,296/2,495
			Low valuations periods				
0.124	**0.040**	**0.045**	**0.056**	-17.97	0.02	-9.14	**-12.79**
(1.72)*	**(1.68)***	**(2.70)****	**(2.82)****	(-1.05)	(0.52)	(-0.54)	**(-1.65)***
[-0.017,0.265]	[-0.007,0.086]	[0.012,0.078]	[0.017,0.095]	[-51.55,15.61]	[-0.057,0.098]	[-42.12,23.84]	[-27.97,2.39]
835	522	852/1,759	852/1,759	841	1,017	858/1,782	858/1,782

Panel D. Instrumental variables: Underwriter rank, log (total assets), float percentage, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies

			Full sample				
0.020	0.035	**0.072**	**0.047**	-5.01	3.50	-3.53	-5.86
(0.61)	(1.53)	**(4.24)****	**(2.09)****	(-0.27)	(0.26)	(-0.36)	(-0.73)
[-0.045,0.086]	[-0.010,0.080]	[0.039,0.105]	[0.003,0.091]	[-40.72,30.70]	[-22.46,29,46]	[-22.90,15.83]	[-21.62,9.91]
2,405	1,959	2,405/5,020	2,405/5,020	2,453	1,997	2,453/5,122	2,453/5,122
			High valuations periods				
0.041	0.018	**0.083**	0.026	6.58	-5.11	7.63	10.72
(0.79)	(0.51)	**(3.03)****	(0.59)	(0.25)	(-1.02)	(0.51)	(0.84)
[-0.061,0.144]	[-0.051,0.087]	[0.029,0.136]	[-0.059,0.111]	[-44.34,57.49]	[-14.88,4.67]	[-21.74,36.99]	[-14.19,35.64]
1,403	1,164	1,268/2,822	1,268/2,434	1,439	1,197	1,296/2,899	1,296/2,495
			Low valuations periods				
0.053	**0.037**	**0.041**	**0.051**	-15.96	-2.6	-12.29	**-13.84**
(1.91)*	**(1.87)***	**(2.68)****	**(2.76)****	(-0.99)	(-0.45)	(-0.96)	**(-1.90)***
[-0.001,0.107]	[-0.002,0.075]	[0.011,0.070]	[0.015,0.088]	[-47.43,15.51]	[-13.99,8.79]	[-37.38,12.81]	[-28.15,0.46]
980	662	996/2,184	996/2,184	992	667	1,008/2,223	1,008/2,223

Table 4.7

Selection bias adjusted first-day returns' and net IPO proceeds' differences between BO-backed and stand-alone IPOs

The table presents selection bias adjusted average net IPO proceeds' and first-day returns' differences between BO-backed and stand-alone IPOs. Each BO-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of BO-backed IPOs matched appear below the average differences. When BO-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. High valuations periods are months when the percentage of IPOs with an offer price above the midpoint of the initial offer range is above median. Low valuations periods are months when this percentage is below median. I take this measure from Jay Ritter's website. Δ(Proceeds) are in millions of dollars and Δ(Returns) in percentage points. ** denotes significance at a 5% level and * at a 10% level.

Δ(Returns)				Δ(Proceeds)			
One-to-one nearest neighbor		One-to-many smoothed weighted		One-to-one nearest neighbor		One-to-many smoothed weighted	
Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Full sample							
-0.008	**-0.027**	0.008	-0.003	**-130.88**	**-32.26**	**-65.71**	**-205.71**
(-0.43)	**(-1.67)***	(0.76)	(-0.26)	**(-3.30)****	**(-3.29)****	**(-3.31)****	**(-3.16)****
[-0.045,0,029]	[-0.058,0.005]	[0.082,0.112]	[-0.022,0.017]	[-208.67,-53.09]	[-51.46,-13.07]	[-104.59,-26.83]	[-333.12,-78.31]
848	697	849/3,581	849/3,581	854	701	855/3,688	855/3,688
High valuations periods							
0.010	0.003	0.010	0.011	**-194.03**	**-61.54**	**-62.57**	**-241.07**
(0.31)	(0.12)	(0.56)	(0.71)	**(-2.70)****	**(-3.93)****	**(-1.96)****	**(-1.90)***
[-0.054,0.074]	[-0.044,0.049]	[-0.024,0.044]	[-0.020,0.043]	[-334.77,-5.33]	[-92.22,-30.87]	[-125.73,-0.60]	[-489.64,7.50]
420	312	423/1,891	423/1,887	425	318	428/1,956	428/1,952
Low valuations periods							
-0.027	**-0.042**	0.003	0.008	**-73.96**	**-49.39**	**-63.37**	**-172.62**
(-1.56)	**(-2.34)****	(0.27)	(0.43)	**(-1.99)****	**(-2.72)****	**(-2.94)****	**(-2.92)****
[-0.062,0.007]	[-0.077,-0.007]	[-0.016,0.022]	[-0.029,0.045]	[-146.99,-0.94]	[-84.97,-13.82]	[-105.65,-21.09]	[-288.62,-56.62]
417	295	422/1,690	422/1,752	418	296	423/1,732	423/1,732
Panel B. Instrumental variables: Underwriter rank, log (total assets), log (equity), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies							
Full sample							
-0.004	**-0.029**	0.011	-0.001	**-156.68**	**-29.51**	**-67.30**	**-190.04**
(-0.21)	**(-1.78)***	(1.00)	(-0.08)	**(-4.09)****	**(-3.01)****	**(-3.43)****	**(-3.67)****
[-0.041,0.033]	[-0.060,0.003]	[-0.010,0.031]	[-0.020,0.018]	[-231.69,-81.67]	[-48.71,-10.31]	[-105.79,-28.81]	[-291.63,-88.45]
848	696	849/3,681	849/3,574	854	701	855/3,681	855/3,681
High valuations periods							
0.017	-0.008	0.008	0.015	**-212.65**	**-62.74**	**-67.18**	**-192.71**
(0.49)	(-0.32)	(0.47)	(0.91)	**(-2.82)****	**(-4.00)****	**(-2.10)****	**(-2.47)****
[0.050,0.084]	[-0.055,0.039]	[-0.027,0.044]	[-0.017,0.047]	[-360.27,-65.04]	[-93.46,-32.03]	[-129.92,-4.44]	[-345.54,-39.87]
424	312	423/1,887	423/1,883	425	318	428/1,952	428/1,948

Table 4.7. Panel B. Cont.

Low valuations periods							
-0.015	**-0.060**	0.009	0.011	**-167.65**	**-34.47**	**-62.70**	**-164.02**
(-0.77)	**(-3.28)****	(0.86)	(0.56)	**(-2.88)****	**(-1.90)***	**(-2.93)****	**(-3.10)****
[-0.052,0.228]	[-0.096,-0.024]	[-0.011,0.028]	[-0.026,0.047]	[-281.72,-53.58]	[-69.95,1.01]	[-104.59,-20.81]	[-267.61,-60.42]
417	295	422/1,687	422/1,687	418	296	423/1,729	423/1,729

Panel C. Instrumental variables: Underwriter rank, log (total assets), default probability, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies

Full sample							
-0.008	-0.009	0.011	0.000	**-170.61**	**-40.86**	**-85.34**	**-195.48**
(-0.37)	(-0.51)	(0.92)	(0.03)	**(-3.73)****	**(-3.53)****	**(-3.89)****	**(-3.69)****
[-0.050,0.0334]	[-0.042,0.024]	[-0.012,0.034]	[-0.022,0.022]	[-260.34,-80.89]	[-63.52,-18.20]	[-128.33,-42.36]	[-299.19,-91.77]
718	570	719/2,792	719/2,792	721	573	722/2,845	722/2,845
High valuations periods							
0.004	0.027	0.019	0.017	**-179.37**	**-63.25**	**-94.40**	**-275.04**
(0.08)	(0.94)	(0.95)	(0.93)	**(-1.81)***	**(-3.60)****	**(-2.53)****	**(-1.92)***
[-0.098,0.107]	[-0.029,0.083]	[-0.021,0.059]	[-0.019,0.054]	[-373.30,14.55]	[-97.68,-28.82]	[-167.52,-21.28]	[-556.32,6.23]
358	247	361/1,531	361/1,527	360	249	363/1,566	363/1,562
Low valuations periods							
-0.012	**-0.046**	0.003	0.019	**-123.66**	**-66.01**	**-71.28**	**-156.69**
(-0.58)	**(-2.16)****	(0.24)	(0.87)	**(-2.10)****	**(-3.05)****	**(-3.12)****	**(-2.88)****
[-0.052,0.282]	[-0.088,-0.004]	[-0.019,0.024)	[-0.023,0.061]	[-239.09,-8.24]	[-108.36,-23.65]	[-116.12,-26.43]	[-263.22,-50.16]
350	220	354/1,261	354/1,261	351	220	355/1,279	355/1,279

Panel D. Instrumental variables: Underwriter rank, log (total assets), float percentage, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies

Full sample							
0.002	-0.013	0.012	0.001	**-206.64**	**-39.12**	**-71.94**	**-205.69**
(0.13)	(-0.78)	(1.14)	(0.07)	**(-4.33)****	**(-3.16)****	**(-3.51)****	**(-3.15)****
[-0.340,0.039]	[-0.045,0.019]	[-0.008,0.032]	[-0.019,0.020]	[-300.16,-113.12]	[-63.40,-14.84]	[-112.11,-31.76]	[-333.79,-77.60)
826	673	827/3,427	827/3,427	830	677	831/3,500	831/3,500
High valuations periods							
0.017	0.013	0.017	0.013	**-174.03**	**-59.11**	**-70.39**	**-225.23**
(0.39)	(0.56)	(0.94)	(0.76)	**(-1.90)***	**(-3.60)****	**(-2.11)****	**(-1.98)****
[-0.068,0.101]	[-0.033,0.060]	[-0.018,0.052]	[-0.020,0.045]	[-353.34,5.28]	[-91.32,-26.90]	[-135.71,-5.06]	[-448.62,-1.85]
410	297	413/1,829	413/1,825	413	301	416/1,874	416/1,870
Low valuations periods							
-0.020	-0.008	0.005	-0.010	**-92.88**	**-32.34**	**-67.20**	**-154.36**
(-1.09)	(-0.43)	(0.50)	(-0.99)	**(-2.27)****	**(-1.68)***	**(-3.04)****	**(-2.90)****
[0.055,0.016]	[-0.045,0.029]	[-0.015,0.025]	[-0.031,0.010]	[-173.19,-12.56]	[-70.15,5.38]	[-110.57,-23.83]	[-258.54,-50.19]
405	282	410/1,598	410/1,598	406	282	411/1,626	411/1,626

Table 4.8

Post-IPO selection bias adjusted default and failure rates' differences between BO-backed and stand-alone IPOs

The table presents selection bias adjusted average default rates' differences between BO-backed and stand-alone IPOs. Each BO-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of BO-backed IPOs matched appear below the average differences. When BO-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. I track companies for five years after the IPO, thus I do not examine IPOs after 2009. I define default as delisting due to bankruptcy or liquidation (CRSP delisting codes 400-490 and 574). I define failure as delisting due to bankruptcy, liquidation or other negative reasons i.e. failure to meet various trading requirements (CRSP delisting codes 400-490, 550-561 and 574-591). ** denotes significance at a 5% level and * at a 10% level.

	One-to-one nearest neighbor		One-to-many smoothed weighted	
	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies				
Default	0.004	0.003	0.119	0.003
	(0.57)	(0.44)	(0.84)	(0.71)
	[-0.010,0.018]	[-0.010,0.016]	[-0.159,0.396]	[-0.006,0.012]
	750	668	750/3,536	750/3,535
Failure	-0.011	-0.002	0.011	-0.020
	(-0.58)	(-0.16)	(1.01)	(-0.87)
	[-0.047,0.025]	[-0.031,0.026]	[-0.010,0.032]	[-0.066,0.026]
	750	668	750/3,536	750/3,535
Panel B. Instrumental variables: Underwriter rank, log (total assets), leverage ratio, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies				
Default	0.000	0.002	0.003	0.003
	(0.00)	(0.31)	(0.35)	(0.60)
	[-0.015,0.015]	[-0.011,0.015]	[-0.012,0.017]	[-0.006,0.011]
	726	644	726/3,347	726/3,347
Failure	-0.029	-0.006	0.016	-0.026
	(-1.42)	(-0.36)	(1.48)	(-1.30)
	[-0.069,0.011]	[-0.036,0.024]	[-0.005,0.037]	[-0.065,0.013]
	726	644	726/3,348	726/3,347
Panel C. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage				
Default	0.012	0.003	0.010	0.006
	(1.20)	(0.46)	(1.64)	(1.53)
	[-0.008,0.032]	[-0.010,0.017]	[-0.0019,0.0216]	[-0.003,0.022]
	497	668	497/2,279	497/2,279
Failure	-0.016	0.002	-0.017	-0.029
	(-0.63)	(-0.14)	(-0.82)	(-0.90)
	[-0.066,0.034]	[-0.033,0.029]	[-0.058,0.0236]	[-0.092,0.034]
	497	668	497/2,279	497/2,279
Panel D. Instrumental variables: Underwriter rank, log (total assets), leverage ratio, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage				
Default	0.010	-0.001	0.001	0.002
	(1.02)	(-0.08)	(0.20)	(0.26)
	[-0.009,0.030]	[-0.015,0.014]	[-0.011,0.014]	[-0.011,0.014]
	490	644	490/2,231	490/2,231
Failure	-0.045	-0.010	-0.010	-0.031
	(-1.52)	(-0.61)	(-0.62)	(-1.10)
	[-0.102,0.013]	[-0.041,0.021]	[-0.041,0.0211]	[-0.085,0.024]
	490	644	490/2,231	490/2,231

Table 4.9

Post-IPO selection bias adjusted default and failure rates' differences between VC-backed and stand-alone IPOs

The table presents selection bias adjusted average default rates' differences between VC-backed and stand-alone IPOs. Each VC-backed IPO is matched with one (nearest neighbor) or many (smoothed weighted) stand-alone IPOs using the propensity score, full covariate, regression-adjusted local linear and inverse probability weighting matching approaches described in the text. I do all matching with replacement and use Abadie and Imbens (2006) standard errors to conduct statistical inference. The t-statistics, 95% confidence intervals and number of VC-backed IPOs matched appear below the average differences. When VC-backed IPOs are matched to many stand-alone IPOs, the total number of observations used in the estimation also appears. I track companies for five years after the IPO, thus I do not examine IPOs after 2009. I define default as delisting due to bankruptcy or liquidation (CRSP delisting codes 400-490 and 574). I define failure as delisting due to bankruptcy, liquidation or other negative reasons i.e. failure to meet various trading requirements (CRSP delisting codes 400-490, 550-561 and 574-591). ** denotes significance at a 5% level and * at a 10% level.

	One-to-one nearest neighbor		One-to-many smoothed weighted	
	Propensity score	Full covariate	Regression-adjusted local linear	Inverse probability weighting
Panel A. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies				
Default	-0.003	0.001	0.004	0.001
	(-0.46)	(0.16)	(0.97)	(0.18)
	[-0.018,0.011]	[-0.011,0.013]	[-0.004,0.011]	[-0.007,0.009]
	2,360	2,035	2,360/5,146	2,360/5,145
Failure	**-0.037**	-0.020	-0.015	**-0.028**
	(-1.85)*	(-1.23)	(-1.39)	**(-2.08)****
	[-0.076,0.002]	[-0.051,0.012]	[-0.037,0.006]	[-0.054,-0.0016]
	2,360	2,035	2,360/5,146	2,360/5,145
Panel B. Instrumental variables: Underwriter rank, log (total assets), return on assets, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies				
Default	-0.007	0.000	0.002	0.001
	(-0.80)	(0.04)	(0.37)	(0.13)
	[-0.024,0.010]	[-0.013,0.014]	[-0.007,0.010]	[-0.009,0.010]
	2,035	1,705	2,035/4,205	2,035/4,205
Failure	**-0.045**	-0.005	-0.016	**-0.053**
	(-2.00)**	(-0.26)	(-1.24)	**(-2.38)****
	[-0.089,-0.001]	[-0.039,0.030]	[-0.040,0.009]	[-0.097,-0.009]
	2,035	1,705	2,035/4,205	2,035/4,205
Panel C. Instrumental variables: Underwriter rank, log (total assets), SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage				
Default	-0.002	-0.001	0.000	0.004
	(-0.17)	(-0.17)	(-0.01)	(0.72)
	[-0.022,0.018]	[-0.013,0.011]	[-0.010,0.010]	[-0.006,0.013]
	1,719	2,035	1,719/3,504	1,719/3,501
Failure	-0.014	-0.018	-0.005	-0.017
	(-0.53)	(-1.15)	(-0.39)	(-1.07)
	[-0.066,0.038]	[-0.049,0.013]	[-0.031,0.021]	[-0.048,0.014]
	1,719	2,035	1,719/3,504	1,719/3,501
Panel D. Instrumental variables: Underwriter rank, log (total assets), return on assets, SIC dummies, headquarter-state dummies, offer year dummies, ln (age), number of total managers, road show success dummies, stock exchange dummies, all-star analyst dummy, syndicate dummy, overallotment percentage				
Default	-0.001	-0.001	-0.003	0.001
	(-0.06)	(-0.18)	(-0.51)	(0.21)
	[-0.021,0.020]	[-0.016,0.013]	[-0.014,0.008]	[-0.010,0.013]
	1,575	1,705	1,575/3,110	1,575/3,110
Failure	-0.020	-0.004	-0.010	**-0.045**
	(-0.70)	(-0.21)	(-0.69)	**(-1.86)***
	[-0.075,0.035]	[-0.038,0.030]	[-0.040,0.091]	[-0.093,0.002]
	1,575	1,705	1,575/3,110	1,575/3,107

Table 4.10

Regressions of default dummy on firm and IPO characteristics

The default models are estimated for 1975-2014 with yearly observations using the multi-period logit technique (Shumway, 2001). The dependent variable is a dummy variable equal to one if the company is delisted due to bankruptcy or liquidation (CRSP delisting codes 400-490 and 574). Financial variables are lagged by one year. NITA is the ratio of net income to total assets. TLTA is the ratio of total liabilities to total assets. To proxy for market conditions at the time of the IPO, I use two measures: The average market first-day return and the logarithm of the number of IPOs in the three months prior the IPO. I construct these measures with data from Jay Ritter's website. Other controls include the ratio of current assets to current liabilities, the natural logarithm of age and interaction terms of industry effects and financial variables (Chava and Jarrow, 2004). Industry fixed effects are based on Chava and Jarrow's (2004) wide industry classifications. Parameter estimates are given first followed by chi-square values in parentheses. ** denotes significance at a 5% level and * at a 10% level.

	(1)		(2)		(3)		(4)	
BO dummy	**-0.614****	**(-2.48)**	**-0.563****	**(-2.24)**	-1.778	(-0.99)	**-0.644****	**(-2.59)**
VC dummy	**-0.416****	**(-2.60)**	**-0.342****	**(-2.04)**	-0.372	(-0.37)	**-0.396****	**(-2.35)**
NITA	**-0.579****	**(-4.05)**	**-0.583****	**(-4.06)**	**-0.581****	**(-4.06)**	**-0.581****	**(-4.02)**
TLTA	**1.755****	**(8.63)**	**1.761****	**(8.66)**	**1.764****	**(8.67)**	**1.762****	**(8.64)**
Log (number of previous IPOs) * BO dummy					0.282	(0.69)		
Log (number of previous IPOs) * VC dummy					0.00635	(0.03)		
Constant	**-7.289****	**(-25.3)**	**-8.763****	**(-16.05)**	**-8.629****	**(-14.40)**	**-7.470****	**(-24.13)**
Other controls		Yes		Yes		Yes		Yes
Industry FE		Yes		Yes		Yes		Yes
Year FE		No		Yes		Yes		Yes
Firm-year observations		65,378		65,378		65,378		65,378
Firms		6,943		6,943		6,943		6,943
Defaulted firms		218		218		218		218
Pseudo R-squared		0.065		0.072		0.072		0.069
Log likelihood		-1,328.2		-1,319.3		-1,319.0		-1,323.4
Wald test		547.95**		580.49**		581.59**		580.39**
Area under Curve		0.77		0.79		0.78		0.78

Table 4.11

Regressions of failure dummy on firm and IPO characteristics

The failure models are estimated for 1975-2014 with yearly observations using the multi-period logit technique (Shumway, 2001). The dependent variable is a dummy variable equal to one if the company is delisted due to bankruptcy, liquidation or other negative reasons i.e. failure to meet various trading requirements (CRSP delisting codes 400-490, 550-561 and 574-591). Financial variables are lagged by one year. NITA is the ratio of net income to total assets. TLTA is the ratio of total liabilities to total assets. To proxy for market conditions at the time of the IPO, I use two measures: The average market first-day return and the logarithm of the number of IPOs in the three months prior the IPO. I construct these measures with data from Jay Ritter's website. Other controls include the ratio of current assets to current liabilities, the natural logarithm of age and interaction terms of industry effects and financial variables (Chava and Jarrow, 2004). Industry fixed effects are based on Chava and Jarrow's (2004) wide industry classifications. Parameter estimates are given first followed by chi-square values in parentheses. ** denotes significance at a 5% level and * at a 10% level.

	(1)		(2)		(3)		(4)	
BO dummy	**-0.664****	**(-6.81)**	**-0.629****	**(-6.36)**	**-1.203***	**(-1.74)**	**-0.665****	**(-6.82)**
VC dummy	**-0.660****	**(-11.12)**	**-0.619****	**(-10.08)**	**-0.884****	**(-2.58)**	**-0.618****	**(-10.05)**
NITA	**-2.197****	**(-3.40)**	**-2.163****	**(-3.37)**	**-2.162****	**(-3.38)**	**-2.221****	**(-3.47)**
TLTA	**1.700****	**(1.98)**	**1.720***	**(1.93)**	**1.712***	**(1.93)**	**1.709****	**(1.96)**
Log (number of previous IPOs * BO dummy					0.135	(0.83)		
Log (number of previous IPOs * VC dummy					0.060	(0.78)		
Constant	**-5.293****	**(-7.67)**	**-6.2341**	**(-8.47)**	**-6.082****	**(-8.14)**	**-5.349****	**(-7.66)**
Other controls		Yes		Yes		Yes		Yes
Industry FE		Yes		Yes		Yes		Yes
Year FE		No		Yes		Yes		Yes
Firm-year observations		65,378		65,378		65,378		65,378
Firms		6,943		6,943		6,943		6,943
Failed firms		1,772		1,772		1,772		1,772
Pseudo R-squared		0.139		0.141		0.145		0.140
Log likelihood		-6,974.5		-6,953.7		-6,953.1		-6,963.3
Wald test		2,317.25**		2,319.19**		2,315.45**		2,328.81**
Area under Curve		0.84		0.84		0.84		0.84

4.7 Figures of Chapter 4

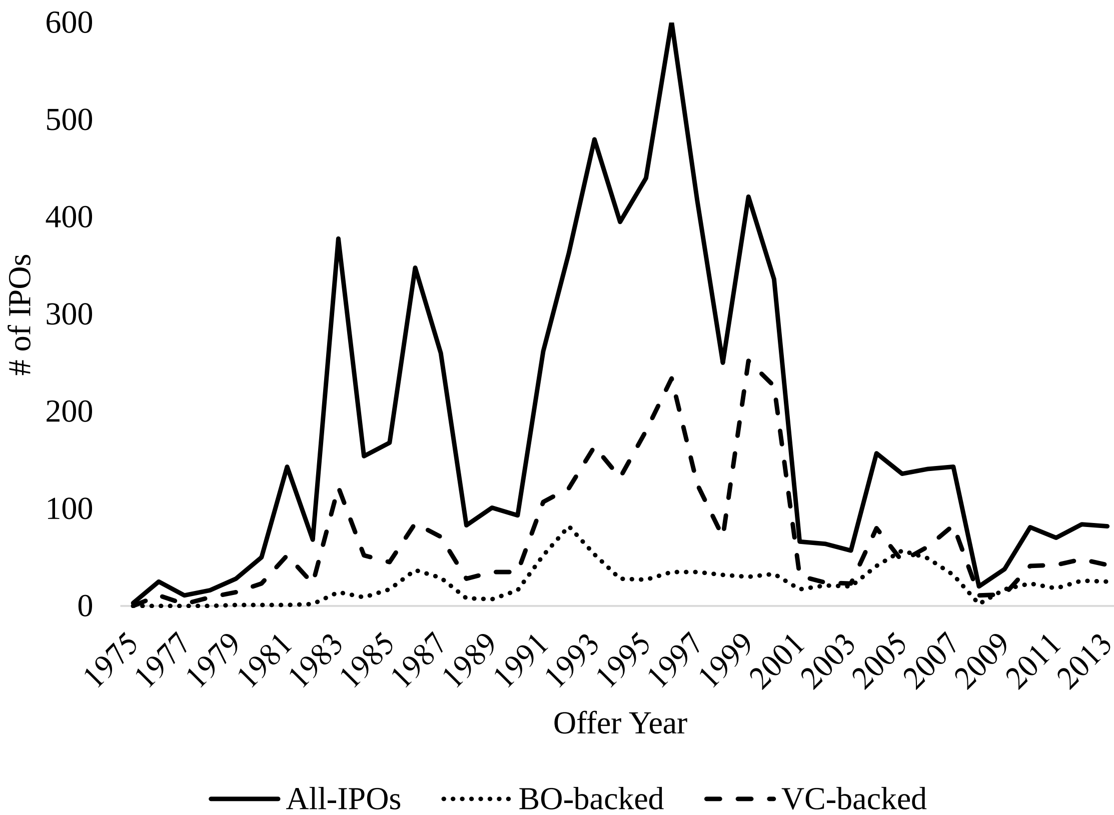

Figure 4.1. IPOs per year. The figure plots the number of all IPOs, BO-backed IPOs and VC-backed IPOs per year for the period 1975 to 2013, excluding those with an offer price below \$5.00 per share, a size below 1.5 million, unit offers, American Depository Receipts (ADRs), closed-end funds, natural resource partnerships, acquisition companies, Real Estate Investment Trusts (REITs), bank and Savings and Loans (S&L) IPOs. I classify growth-capital backed IPOs as VC-backed.

Conclusions

The three papers of this thesis explore default and credit performance determinants in private firms (SMEs), listed firms (especially mid and small-caps), and companies that are in the transition phase between private and public ownership (IPOs) respectively. The work has several implications for banks, investors and policy makers by (i) uncovering regional corporate default risk vulnerabilities, (ii) explaining the observed anomalous pricing of default risk in the stock market, and (iii) exploring the impact of alternative investment funds such as private equity and venture capital on the default risk of their portfolio companies.

The first paper uses macro-economic variables to compliment firm-specific data, enhancing the models' forecasting ability, especially at the aggregate default incidence level for different sovereign environments in Europe. So, while fundamental, firm-specific variables show important stable and robust levels of accuracy across countries, the aggregate level of distress is considerably enhanced by adding macro-determinants and cross-country data. The paper validates the superiority of models that incorporate macroeconomic dependencies, suggested by previous research, also in the case of SMEs. Specifically, macro variables differ among European regions based on region-specific conditions and characteristics. Since our regional distress models always perform better than a generic model estimated for the regional sub-samples, we conclude that their use can lead to performance improvements in the risk management of international SME portfolios.

The second paper tackles a thorny and controversial issue in Finance involving risk and return expectations versus actual results. We provide an explanation for the anomalous pattern based on an argument that rests on the impact of systematic default risk (SDR) on returns of individual firms. We show that the physical PD is usually a poor measure of exposure to aggregate default risk because stocks in the highest PD quintile have relatively low SDR exposure. Investors indeed require a premium to hold stocks that are riskier when aggregate default risk is higher but they do not require compensation to hold stocks with high firm-specific risk because these stocks are a source of portfolio risk diversification. Our results suggest that riskier stocks, as measured by the physical PDs, will tend to under-perform because they have, on average, lower exposure to aggregate default risk. Their riskiness is mostly idiosyncratic and can be diversified away. On the contrary, it is the systematic component of default risk, measured by the SDR betas, that requires a return premium.

The third paper involves assessing the impact of PE-backed firms which tap the IPO market as an exit strategy and compares these IPOs with non-PE backed IPOs. The different incentives of PE investors and managers can strongly impact subsequent performance and default risk. These professional insiders may be more capable of taking advantage of information asymmetries compared to insiders of stand-alone companies. But they also have more reputational capital at stake, a factor which tends to be known by the market. We find evidence against the hypothesis that PE sponsors "cheat" the market . The financial situation of both BO and VC-backed companies in the pre-IPO year, as measured by their default risk, is similar to that of their peers. Moreover, PE sponsors do not target their IPOs in hot periods any more than do managers of stand-alone companies. They also are not more prone to rush their companies into premature IPOs and do not inflate valuations. Finally, PE-backed companies do not default more often post-IPO. This is evidence that PE sponsors are not more likely to seek to sell firms with poor prospects ("unload lemons") in the IPO market.

References

Abadie, Alberto, and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects". Econometrica 74(1), 235-267.

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

Aggarwal, Rajesh K., Laurie Krigman, and Kent L.Womack. 2002. "Strategic IPO Underpricing, Information Momentum, and Lockup Expiration Selling". Journal of Financial Economics 66(1), 105-137.

Alti, Aydogan. 2005. "IPO Market Timing". The Review of Financial Studies 10(3), 1105-1138.

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

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

Altman, Edward I., Małgorzata Iwanicz-Drozdowska, Erkki K. Laitinen, Arto Suvas. 2014 "Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model". Working Paper.

Altman, Edward I., and Herbert Rijken. 2012. "Toward A Bottom-Up Approach to Assessing Sovereign Default Risk: An Update". Working Paper.

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

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

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

Andrade, Gregor, and Steven N. Kaplan. 1998. "How Costly is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions that Became Distressed". The Journal of Finance 53(5), 1443-1493.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006. "The Cross-Section of Volatility and Expected Returns". The Journal of Finance 61(1), 259-299.

Anginer, Deniz, and Çelim Yıldızhan. 2014. "Is there a Distress Risk Anomaly? Pricing of Systematic Default Risk in the Cross Section of Equity Returns". Working Paper.

Aretz, Kevin, Chris Florackis, and Alexandros Kostakis. 2014. "Do Stock Returns Really Decrease with Default Risk? New International Evidence". Working paper.

Avramov, Doron, Scott Cederburg, and Satadru Hore. 2011. "Cross-Sectional Asset Pricing Puzzles: A Long-Run Perspective". Working Paper.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov. 2009. “Credit Ratings and the Cross-Section of Stock Returns”. Journal of Financial Markets 12(3), 469-499.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov. 2013. “Anomalies and Financial Distress”. Journal of Financial Economics 108(1), 139-159.

Ball, Eric, Hsin Hui Chiu, and Richard Smith. 2011. “Can VCs Time the Market? An Analysis of Exit Choice for Venture-backed Firms?”. The Review of Financial Studies 24(9), 3105-3138.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron. 2014. “Volatility, the Macroeconomy, and Asset Prices”. The Journal of Finance 69(6), 2471-2511.

Barinov, Alexander. 2012. “Aggregate Volatility Risk: Explaining the Small Growth Anomaly and the New Issues Puzzle”. Journal of Corporate Finance 18(4), 763-781.

Barry, Christopher B., Chris J. Muscarella, John W. Peavy III, and Michael R. Vetsuypens. 1990. “The Role of Venture Capital in the Creation of Public Companies: Evidence from the Going-Public Process”. Journal of Financial Economics 27(2), 447–71.

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

Bauer, Julian, Vineet Agarwal. 2014 “Distress Risk and Stock Returns: The Neglected Profitability Effect”. Working Paper.

Bayar, Onur, and Thomas J. Chemmanur. 2011. “IPOs versus Acquisitions and the Valuation Premium Puzzle: A Theory of Exit Choice by Entrepreneurs and Venture Capitalists”. Journal of Financial and Quantitative Analysis 46(6), 1755-1793.

Bayar, Onur, and Thomas J. Chemmanur. 2012. “What Drives the Valuation Premium in IPOs versus Acquisitions? An Empirical Analysis”. Journal of Corporate Finance 18(3), 451-475.

Becchetti, Leonardo, and Jaime Sierra. 2003. “Bankruptcy Risk and Productive Efficiency in Manufacturing Firms.” Journal of Banking and Finance 27(11), 2099-2120.

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

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

Black, Fischer, and Myron Scholes. 1973. “The Pricing of Options and Corporate Liabilities”. Journal of Political Economy 81(3), 637-654.

Bradley, J. Daniel, Bradford D. Jordan, and Jay R. Ritter. 2008. “Analyst Behavior Following IPOs: The “Bubble Period” Evidence. The Review of Financial Studies 21(1), 101-133.

Brav, Alon, and Paul A. Gompers. 1997. “Myth or Reality? The Long-Run Underperformance of Initial Public Offerings: Evidence from Venture and Nonventure Capital-Backed Companies”. The Journal of Finance 52(5), 1791–821.

Bonfim, Diana. 2009. “Credit Risk Drivers: Evaluating the Contribution of Firm Level Information and of Macroeconomic Dynamics”. Journal of Banking and Finance 33(2), 281-299.

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

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

Cao, Jerry X. 2011. “IPO Timing, Buyout Sponsors’ Exit Strategies and Firm Performance of RLBOs”. Journal of Financial and Quantitative Analysis 46(4), 1001-1024.

Cao, Jerry X., and Josh Lerner. 2009. “The Performance of Reverse Leveraged Buyouts”. Journal of Financial Economics 91(2), 139-157.

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

Chan-Lau, J. A. 2006. “Is Systematic Default Risk Priced in Equity Returns? A Cross-sectional Analysis using Credit Derivatives Prices”. IMF working paper, 148.

Chava, Sudheer, Amiyatosh Purnanandam. 2010. “Is Default Risk Negatively Related to Stock Returns?”. The Review of Financial Studies 23(6), 2523-2559.

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

Chemmanur, Thomas J., and Paolo Fulghieri. 1990. “A Theory of the Going-Public Decision”. The Review of Financial Studies 12(2), 249-279.

Chen, Long, and Lu Zhang. 2010. “A Better Three-Factor Model That Explains More Anomalies”. The Journal of Finance 65(2), 563-594.

Chira, Inga, Jeff Madura, and Ariel M. Vialeb. 2013. “Bank Exposure to Market Fear”. Journal of Financial Stability 9(4), 451-459.

Cliff, Michael T., and David J. Denis. 2005. “Do Initial Public Offering Firms Purchase Analyst Coverage with Underpricing?”. The Journal of Finance 59(6), 2871-2901.

Collin-Dufresne, Pierre, Robert S. Goldstein, J. Spenser Martin. 2001. “The Determinants of Credit Spread Changes”. The Journal of Finance 56(6), 2177-2207.

Conrad, Jennifer, Nishad Kapadia, and Yuhang Xing. 2012 “What Explains the Distress Risk Puzzle: Death or Glory?”. Working paper.

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

Creditreform. 2012 Insolvencies in Europe. Surveys by the Creditreform Economic Research Unit, available at http://www.creditreform.com/en/press-and-downloads/analysis.html.

Da, Zhi, and Pengjie Gao. 2010. “Clientele Change, Liquidity Shock, and the Return on Financially Distressed Stocks”. Journal of Financial and Quantitative Analysis 45(1), 27-48.

Davydenko, Sergei A. 2012. “When Do Firms Default? A Study of the Default Boundary”. Working Paper.

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

Degeorge, Francois, and Richard Zeckhauser. 1993. "The Reverse LBO Decision and Firm Performance: Theory and Evidence". The Journal of Finance 48(4), 1323-1348.

Dichev, Ilia D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?". The Journal of Finance 53(3), 1131-1147.

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

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

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

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

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

Fang, Lily, and Ayako Yasuda. 2009. "The Effectiveness of Reputation as a Disciplinary Mechanism in Sell-Side Research". The Review of Financial Studies 22(9), 3735-3777.

Field, Laura C., and Gordon Hanka. 2001. "The Expiration of IPO Share Lockups". The Journal of Finance 56(2), 471-500.

Field, Laura C., and Jonathan Karpoff. 2002. "Takeover Defenses of IPO Firms". The Journal of Finance 57(5), 1857-1889.

Friedwald, Nils, Christian Wagner, and Josef Zechner. 2014. "The Cross-Section of Credit Risk Premia and Equity Returns". The Journal of Finance 69(6), 2419-2469..

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

Gao, Pengjie, Christopher A. Parsons, and Jainfeng Shen. 2013. "The Global Relation Between Financial Distress and Equity Returns". Working paper.

Garlappi, Lorenzo, and Hong Yan. 2011. "Financial Distress and the Cross-Section of Equity-Returns". The Journal of Finance 66(3), 789-822.

Garlappi, Lorenzo, Tao Shu, and Hong Yan. 2008. "Default Risk, Shareholder Advantage, and Stock Returns". The Review of Financial Studies 21(6), 2743-2778.

George Thomas J., and Chuan-Yang Hwang. 2010. "A Resolution of the Distress Risk and Leverage Puzzles in the Cross Section of Stock Returns". Journal of Financial Economics 96(1), 56-79.

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

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

Gompers, Paul A. 1996. “Grandstanding in the Venture Capital Industry”. Journal of Financial Economics 42(1), 133-156.

Gormley, Torr A., and David A. Matsa. 2014. “Common Errors: How to (and Not to) Control for Unobserved Heterogeneity”. Review of Financial Studies 27(2), 617-61.

Grammatikos, Theoharry, and Robert Vermeulen. 2012. “Transmission of the Financial and Sovereign Debt Crises to the EMU: Stock Prices, CDS Spreads and Exchange Rates”. Journal of International Money and Finance 31(3), 517-533.

Griffin, John M., and Michael L. Lemmon. 2002. “Book-to-Market Equity, Distress Risk and Stock Returns”. The Journal of Finance 57(5), 2317-2336.

Grunert Jens, Lars Norden, and Martin Weber. 2005. “The Role of Non-Financial Factors in Internal Credit Ratings.” Journal of Banking and Finance 29(2), 509-531.

Harford, Jarrad, and Adam Kolasinski. 2014. “Do Private Equity Returns Result from Wealth Transfers and Short-Termism? Evidence from a Comprehensive Sample of Large Buyouts”. Management Science 60(4), 888-902.

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

Heckman, James. 1997. “Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations”. Journal of Human Resources 32(3), 441-462.

Hilscher, Jens, and Mungo Wilson. 2013. “Credit Ratings and Credit Risk: Is One Measure Enough?”. Working Paper.

Hogan, Karen M., Gerard T. Olson, and Richard J. Kish. 2001. "A Comparison of Reverse Leveraged Buyouts and Original Initial Public Offers: Factors Impacting Their Issuance in the IPO Market". Financial Review 36(3), 1-18.

Holthausen, Robert W., and David F. Larcker. 1996. “The Financial Performance of Reverse Leveraged Buyouts”. Journal of Financial Economics 42(3), 293-332.

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

Hotchkiss, Edith S., David C. Smith, and Per Stomberg. 2014. “Private Equity and the Resolution of Financial Distress”. Working Paper.

Ibbotson, Roger G., Jody L. Sindelar, and Jay R. Ritter. 1994. “The Market’s Problems with the Pricing of Initial Public Offerings”. Journal of Applied Corporate Finance 7(1), 66-74.

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

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

Jain, Bharat A., and Omesh Kini. 2000. “Does the Presence of Venture Capitalists Improve the Survival Profile of IPO Firms?”. Journal of Business Finance and Accounting 27(9-10), 1139-1176.

Jaumotte, Florence, and Piyaporn Sodsriwiboon. 2010. “Current Account Imbalances in the Southern Euro Area”. IMF Working Paper WP/10/139.

Kapadia, Nishad. 2011. “Tracking down Distress Risk”. Journal of Financial Economics 102(1), 167-182.

Kaplan, Steven N., and Jeremy C. Stein. 1993. “The Evolution of Buyout Pricing and Financial Structure in the 1980s”. The Quarterly Journal of Economics 108(2), 313-357.

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

Krishnan, C. N. V., Vladimir I. Ivanov, Ronald W. Masulis, and Ajai K. Singh. 2011. “Venture Capital Reputation, Post- IPO Performance and Corporate Governance”. Journal of Financial and Quantitative Analysis 46(5), 1295-1333.

Laekholm-Jensen, Thais, David Lando, Mamdouh Medhat. 2013. “Cyclicality and Firm Size in Private Firm Defaults”. Working Paper.

Lee, Peggy M., and Sunil Wahal. 2004. “Grandstanding, Certification and the Underpricing of Venture Capital Backed IPOs”. Journal of Financial Economics 73(2), 375-407.

Lerner, Joshua. 1994. “Venture Capitalists and the Decision to Go Public”. Journal of Financial Economics 35(3), 293-316.

Lewellen, Jonathan. 1999. “The Time-Series Relations among Expected Returns, Risk and Book-to-Market”. Journal of Financial Economics 54(1), 5-43.

Liu, Xiaoding, and Jay R. Ritter. 2011. “Local Underwriter Oligopolies and IPO Underpricing”. Journal of Financial Economics 102(3), 579-601.

Loughran, Tim, and Jay R. Ritter. 2004. “Why Has IPO Underpricing Changed Over Time?”. Financial Management 33(3), 5-37.

Megginson, William L., and Kathleen A. Weiss. 1991. “Venture Capitalist Certification in Initial Public Offerings”. The Journal of Finance 46(3), 879-903.

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

Metrick, Andrew, and Ayako Yasuda. 2011. “Venture Capital and Other Private Equity: A Survey”. European Financial Management 17(4), 619-654.

Miller, Edward M. 1977. “Risk, Uncertainty, and Divergence of Opinion”. The Journal of Finance 32(4), 1151-1168.

Muscarella, Chris J., and Michael R. Vetsuypens. 1990. "Efficiency and Organizational Structure: A Study of Reverse LBOs". The Journal of Finance 45(5), 1389-413.

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

Neus, Warner, and Uwe Walz. 2005. "Exit Timing of Venture Capitalists in the Course of an Initial Public Offering". Journal of Financial Intermediation 14(2), 253-277.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". Econometrica 55(3), 703-708.

Nielsen, Caren Yinxia G. 2013. "Is Default Risk Priced in Equity Returns?". Working Paper.

OECD. 2005. OECD SME and Entrepreneurship Outlook 2005.

Ohlson, James A. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy". Journal of Accounting Research 18(1), 109-131.

Ozdagli, Ali K. 2013. "Distressed, but not risky: Reconciling the empirical relationship between financial distress, market-based risk indicators, and stock returns (and more)". Working paper.

Pan, Jun, and Kenneth J. Singleton. 2008. "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads". The Journal of Finance 63(5), 2345-2384.

Pastor, Lubos, and Pietro Veronezi. 2005. "Rational IPO Waves". The Journal of Finance, 60(4). 1713-1757.

Perego, Erika R., Wessel N. Vermeulen. 2013. "Macroeconomic Determinants of European Stock and Government Bond Correlations: A Tale of Two Regions". Working Paper.

Puri, Manju, and Rebecca Zarutskie. 2012. "On the Life Cycle Dynamics of Venture-Capital-and Non-venture-Capital-Financed Firms". The Journal of Finance 67(6), 2247-2293.

Ritter, Jay R., and Ivo Welch. 2002. "A Review of IPO Activity, Pricing, and Allocations". The Journal of Finance 57(4), 1795-1828.

Rosenbaum, Paul. R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects". Biometrika 70(1), 41-55.

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

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". Journal of Educational Psychology 66(5), 688-701.

Rubin, Donald B. 1977. "Assignment to Treatment Group on the Basis of a Covariate". Journal of Educational Statistics 2(1), 1-26.

Schaefer, Stephen M., and Ilya A. Strebulaev. 2008. "Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds". Journal of Financial Economics 90(1), 1-19.

Schultz, Paul. 2003. "Pseudo Market Timing and The Long-Run Underperformance of IPOs". The Journal of Finance 58(2), 483-518.

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

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

Tykvova, Tereza, and Mariela Borell. 2012. “Do Private Equity Owners Increase Risk of Financial Distress and Bankruptcy?”. Journal of Corporate Finance 18(1), 138-150.

Vassalou, Maria, and Yuhang Xing. 2004. “Default Risk in Equity Returns”. The Journal of Finance 59(2), 831-868.

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

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

Wilson, Nick, and Mike Wright. 2013. “Private Equity, Buyouts and Insolvency Risk”. Journal of Business Finance and Accounting 40(7-8), 949-990.

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

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

Yung, Chris, Gonum Colak, and Wei Wang. 2008. Cycles in the IPO Market. Journal of Financial Economics 89(1), 192-208.

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

Vita

Dimitra Michala

Degrees

2015	Ph.D. in Finance, University of Luxembourg, Luxembourg School of Finance
2014	Visiting Scholar, New York University, Stern School of Business
2009	M.B.A. in Financial Engineering, Athens University of Economics and Business
2006	B.Sc. in Finance and Banking, University of Piraeus

Positions Held

2015-to date	Economist, Commission de Surveillance du Secteur Financier
2011-2015	Researcher and Instructor, University of Luxembourg
2009-2011	Junior Auditor, European Court of Auditors, Audit Methodology Unit
2009-2009	Intern Analyst, Grant Thornton, Transactions Advisory and Valuation Unit
2007-2008	Business Credit Officer, Bank of Cyprus