Drivers of Corporate Bankruptcy and Default - Study Concerning Portuguese Companies

JOSÉ CARLOS MORAIS

ADVISOR – Diana Bonfim

Dissertation submitted in partial fulfillment of requirements for the degree of Master in Finance, at the Universidade Católica Portuguesa, date: 2013

ABSTRACT

My purpose with this dissertation is to investigate the relevance of financial and economic variables as predictors of corporate bankruptcy (default) in Portugal. Understanding the underlying issues behind corporate distress and failure is crucial for several areas particularly: granting (or not) a bank loan, managing credit risk and pricing debt. The present financial and economic crisis stimulated a growing interest on matters related to corporate default, ratings and scoring models. In this dissertation, I apply, to a sample of medium-large Portuguese companies, updated statistical models for identifying the fundamental variables that lead to bankruptcy (and default), and if they are mainly idiosyncratic and / or systematic, affecting simultaneously all firms. I also fit hazard model for studying hazard rates and survival-time, trying to disclose which variables can shield firms from a hazard event, in this case, bankruptcy. The sample was provided by COFACE, a known credit scoring company and included full annual financial statements, from years 2006 to 2011, adding additional requested information, namely, about bankruptcies and legal actions against firms due to payment default. My findings support the conclusion that some financial and economic variables do influence bankruptcy and default probability together with survival-time, which is perfectly anticipated by corporate finance analysis theory and practice. Nevertheless, I reached some conclusions that contradict others from similar studies.

PREFACE

My interest in proceeding with a presentation of a master thesis dissertation was prior to my presence in the 6th Master in Finance program edition in Católica Lisbon SBE; this interest solidified during the program, ending in a firm decision of presenting a work of the kind. My main area of interest is corporate finance and corporate risk so, I would intend to develop a thesis related to those matters and, if possible, applied to Portuguese real examples. This possibility arose during dissertation seminar lead by Dra. Diana Bonfim, particularly, regarding corporate risk analysis. Matters related to Probability of Default (PD) and Credit Risk and scoring, namely for corporate sector, are of growing interests and relevance for financial sector.

The issue was now to seek a real sample of Portuguese enterprises in order to be able to apply updated methodologies for analyzing corporate risk. This sample was provided by COFACE®, the credit scoring and insurance company, which gave me access to their extensive database base that holds full financial statements from, practically, all enterprises in Portugal. Besides full financial statements data, COFACE database comprises several types of information's, relevant for analyzing corporate risk.

This dissertation was only possible with the help of Dra. **Diana Bonfim** that guided me trough all the process, pointing where could I find and study the best practices and theories in this field and, also, in the use of the statistical software package **STATA®,** where all analysis and calculations were performed. My final acknowledgement goes to **COFACE®**. Without their help and contribution, allowing the access to their data, I would be unable to present this thesis. This is part of **COFACE®** corporate social responsibility police; the availability to contribute to scientific and academic research on matters related to financial risk.

It was for me a pleasurable, challenging and demanding experience developing this thesis where I was able to acquire knowledge and practices that were unknown to me, particularly, in micro-econometrics, and corporate risk analysis, that, I am sure, will be strongly useful in my future professional life.

CONTENTS

INDEX of Tables

1. INTRODUCTION

Credit risk has been historically an area of close attention, not only in banks and financial sector, but also, at corporate level. Banks need to forecast precisely the default risk on their assets in order to adequately protect their array of financers – deposit holders, bond holders, other banks, equity holders, etc – by accurately managing and pricing their operations or eventually deciding not to go forward.

In enterprises, default risk is present, particularly, in accounts receivable. The possibility of a costumer defaulting is an issue that needs to be managed when setting sales terms. Business margins and profitability can be severely jeopardized by costumers who do not comply with payment terms or, in a worst-case-scenario, fully or partially default – particularly if the costumer goes bankrupt.

With the recent international financial and economic turmoil, which triggered severe economic recessions in several developed countries, managing and studying these matters gained growing attention.

The generic matter that needs to be answered is: What are the main triggers of company bankruptcy and default? **Particularly, regarding Portuguese companies, what are the main drivers resulting in default, and/or bankruptcy?** These are the decisive questions in order to manage credit risk. Specifically in the banking sector, credit risk is a decisive issue to handle, given the weight corporate loans have on banks assets. The ongoing financial crisis facing the financial services industry, created challenges for future credit risk modeling.

Following Basel II (June 2004) agreement more sophisticated capital adequacy rules were issued for banks worldwide. Under the Basel II framework, a new methodology to compute minimum standards for capital adequacy was issued, and, all banking authorities present in the committee implemented the adoption in their home countries. Basel II presents an important development regarding banks' own models of **credit risk measurement**, which have become more sophisticated **tools for assessing risk for capital requirements**. The Basel II framework is fully compiled in the 2006 document "International Convergence of Capital Measurement and Capital Standards". Banks are ever more encouraged to move to more sophisticated models to estimate Probabilities of Default (PD's), essentially founded in **econometric scoring models**. However, it should not be excluded from these estimations "human judgment" and relevant variables external to scoring models that need to be taken in consideration (ex. Corporate Governance). In many cases, Basel II was implemented immediately before the crisis. Internal credit risk models were built using data from a relatively tranquil period, usually labeled as the "great moderation". Given this, the analysis of determinants of PD's during the crises may provide new insights on this issue.

From this need appears the main interest and motivation for this dissertation: **to apply most updated theories and practices in credit risk models and estimation of Probabilities of Default (PD) to a selected and representative sample of non-financial Portuguese companies (publicly listed and non listed), trying to identify which factors can predict company default or bankruptcy.**

This dissertation is based on similar works done in Portugal as in several other countries. The general purpose of those papers is to identify variables that influence corporate default and that somehow allow the forecast of similar events. One difference in this thesis is related to the economic turmoil Portugal is facing; the sample where this study is based has information's for two years before the crisis period started, and four years after 2008, when events were triggered. Conclusions may be disclosed regarding the influence this economic turmoil certainly has on corporate defaults and bankruptcies.

The research will start by introducing best practices and theories for corporate default prediction, describe the dataset, summarize its main characteristics and proceed with the econometric modeling and analyses. The thesis will end with a generic description of the main conclusions disclosed during the study.

2. LITERATURE REVIEW AND MAIN DEVELOPMENTS

Financial distress prediction models, for both default and bankruptcy, have been extensively used for some decades. Different kind of models and papers regarding these matters have been issued, being some focused on large listed public companies, others on non-listed private companies; concentrated solely on firm accounting information, others exclusively on public information, still others, mixing both indicators, and also, adding economic variables.

2.1 Eduard Altman

Altman (1968) seminal work is consensually considered a reference and historical starter of scoring models for corporations, still very popular presently. He started with, what was at the time, the main tool for analyzing corporations – ratio analysis - and, due to the insufficiencies on this method he tried to take a step forward improving the quality of the tool as an analytical technique. The prediction of corporate bankruptcy was used as an illustrative case. The statistical technique chosen for the task was multiple discriminant analysis (MDA). Although not as popular as regression analysis was, at the time, it was increasingly being used in several science fields.

The sample consisted in 66 manufacturing companies, divided in two 33 companies groups: non-bankrupt and bankrupt. Historical data ranged from 1946 to 1965, and tests were made using financial statements one year prior to bankruptcy. A set of variables (financial ratios) was tested and carefully selected for the purpose of bankruptcy prediction; each variable was weighted for contribution for the model and, a final function was selected as being the one achieving best results for the purpose of the task, a Z-score as it was called onwards. A score ranking resulted from the analysis from -4 (near bankruptcy) to 8 (probability of bankruptcy near 0). **The model delivered an accurate forecast up to two years before bankruptcy. Accuracy diminishes substantially as lead time increases.** The Z-score model is summarized in the following expression:

$$\boldsymbol{Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5}$$

Where: X_1 = working capital/total assets ratio

X_2 = retained earnings/total assets ratio

X_3 = earnings before interests and taxes/Total assets ratio

X_4 = market value of equity/book value of total liabilities ratio

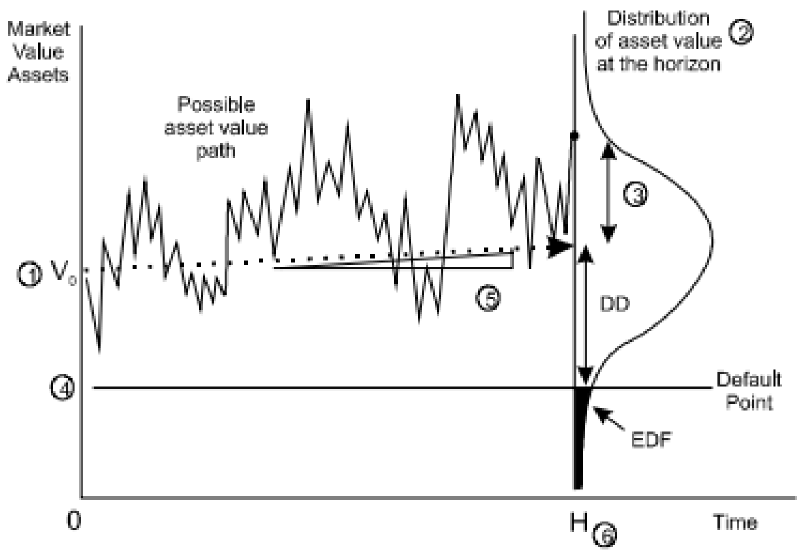
X_5 = sales/total assets ratio

2.2 Robert C. Merton

An additional contribution was given by Merton, Robert C. (1974). The purpose of the work was to apply option pricing theory to valuation of risky bonds. Until then, there had been no systematic theory of pricing bonds when there is a significant probability of default. It develops from the Black F. and Scholes M. (1973) general equilibrium theory of option pricing, particularly attractive since, is a function of observable variables, and, can be subjected to empirical tests. Some assumptions are needed in order to develop the model: There is a "perfect market" environment; the M. Miller and F. Modigliani. (1958) irrelevance theorem holds; the Term-Structure of riskless interest rate is known with certainty and flat; the dynamics for the value of a firm can be described by a diffusion-type stochastic process.

2.2.1 Further developments on Merton's work

One of the developments of Robert C. Merton model is the estimation of default prediction models that produce default predictions for companies and banks that have their equity publicly traded, as the well known **Moody's KMV model**. This model considers the loan repayment incentives from the view point of the borrowing firm's equity holders - for further details see Saunders A. and Allen L. (2010) and Crosbie P. and Kocagil A. (2003) -. The market-value position of equity holders in a leveraged firm can be viewed as isomorphic to holding a call option on the assets of the firm; in limited liability firms, equity holders will keep the residual market value of firm's assets after repaying loan at maturity. The larger the market value of assets, the larger the payoff of equity holders; Otherwise, if the market value of firm's assets falls below the loan amount at maturity, the company will be insolvent and equity holders have the incentive to turn firm´s assets to the bank. So, equity holders have limited downside risk. Moody's KMV model calculates a probability of default measure, or **Expected Default Frequency (EDF™)**; essentially, a distance to default – the amount by which assets market value would fall for company to default - measure (DD) is obtained by a ratio between the difference from market value of assets and loan amount divided by assets volatility (measured by the standard deviation). Assuming assets values are normally distributed, allows to calculate the probability of assets value to enter default region. From Crosbie P. and Kocagil A., (2003): "*there are six variables that determine the default probability of a firm over some horizon, from now until H (see below figure1): (1) the current assets value; (2) the distribution of assets value at time H; (3) the volatility of future assets value at time H; (4) the level of the default point, the book value of liabilities; (5) the expected rate of growth in the assets value over the horizon; (6) the length of the horizon, H.*"



2.2.2 Strengths and Weakenesses

There are a number of strengths in using option pricing theory to model default prediction: (1) it can be applied to any public company; (2) not being based on accounting information but on stock market date is forward looking; (3) it is grounded in solid economic analysis. The main weaknesses: (1) the normality assumption on asset returns; (2) for private firms it can only be applied if some similarities are found between accounting data and other observable characteristics of the borrower; (3) it does not distinguish between different types of debt; (4) the Merton model is static since it assumes a constant debt structure, thus it cannot capture the behavior of those firms who seek an optimal debt-to-equity ratio.

Bharath T. S. and Shumway T. (2004) examine the accuracy of KMV model default forecasting against a similar but simpler alternative model. They used data from all public non-financial firms in NYSE, AMEX and NASDAQ from 1980 to 2003. Defaults were collected from Altman default database and Moody's. They concluded that the KMV-Merton model does not appear to be a sufficient statistic for default; the simpler model proposed which captures some inputs of KMV-Merton model, performs surprisingly well.

2.3 Reduced Form Models

Models that use information's contained in equity prices are called **Structural Models** - for details on additional default probabilities models see Saunders A. and Allen L. (2010), chapters 5 and 6. **Reduced Form Models** use other securities prices to achieve the same goal. These models utilize the information embedded in risky bonds yield decomposed into risk-free rate - the return obtained in a risk-free asset - plus a risk premium, using this decomposition to calculate default probabilities. In a market where investors behave in a risk-neutral fashion, the price of any asset could be calculated by discounting all forecasted cash-flows by the risk-free rate; this relationship can be used to calculate the default risk-premium of any asset. From the procedure of these models comes their major setback: the reliance on noisy bond price data. The difference in prices from risky-bonds and the equivalent risk-free may result, not only, from credit risk, but also from other issues: taxes, liquidity premium, data/pricing errors. Several papers conclude that **reduced form models are the state of the art in default risk estimation** - See Campbel l, J., J. Hilscher, and J. Szilagyi (2008) -.

2.3.1 KRIS™

An example of a reduced form model in use is the KRIS™: Kamakura's Risk Information Services; from their web:

> *"Founded in 1990 by Dr. Donald R. van Deventer, Kamakura Corporation is the world's leading provider of risk management information, risk management software and risk management consulting. Kamakura's executive team represents a broad and diverse cross-section of in-depth experience in economics, financial management, information*

technology, credit modeling, risk assessment, accounting, business administration, higher education, banking and regulatory oversight." From: http://www.kamakuraco.com/ .

Estimates of probability of default (PD) in KRIS™ model incorporate firm-specific information, industry information, economic environment, and macroeconomic factors. The first version of KRIS model used information from US companies from 1962 to 1990; five explanatory variables were used in the model: (1) Return on Assets; (2) Leverage; (3); Relative size in NYSE and AMEX (4); Excess Return (5); Monthly equity volatility. In March 2009, Kamakura upgraded the KRIS model to incorporate 40 key macroeconomic risk factors into the estimation of default probabilities for more than 20.000 public firms worldwide. Chava S. and J. A. Jarrow (2004) find that public firm model including firm-specific accounting variables has a 91,98% accuracy rate (compares to 90% accuracy from KMV Moody's model).

2.4 Other Models

Other Credit Risk Models, more established, which have been applied for several decades are: (1) Credit Scoring Models, (2) Mortality Rate Systems, (3) Neural Network Systems. Further details in Saunders A. and Allen L. (2010), chapter 6.

2.4.1 Credit Scoring Models

Credit Scoring Models are extensively used in several credit analyses, from consumer credit to corporate loans. These models try to identify certain key factors that determine the probability of default weighting them into a score that can be used as a probability of default, and scaling borrowers according to their credit quality. There are four main methodological forms of multivariate credit scoring models: (1) linear probability model, (2) the logit model, (3) the probit model, and (4) the discriminant analysis model. One of the most popular credit scoring model is the priviously mentioned Altman Z-Score. Over the time the initial model suffered several developments and, presently, Z'' (or Z-double prime) is the most updated version. Once calculated, the score can be mapped into an equivalent agency rating. Some setbacks need to be mentioned for this model: (1) the model is linear and the path to bankruptcy may be nonlinear; (2) the *Z*-score is based on accounting ratios and accounting data is disclosed at discrete intervals and are generally based on historical principles, thus being questionable whether such models are able to identify firms whose conditions are rapidly worsening.

2.4.2 Mortality Rate Systems

Mortality Rate Systems are based on historical rates of default and can be used for default prediction one year ahead, marginal mortality rates (MMR) or multiyear, cumulative mortality rates (CMR). For each rating grade, the analyst will pick a sample of issued years, and for each year, will sum the total value of bonds defaulting dividing it by the total amount of bonds issued, on the particular grade he is analyzing; MMR will be the weighted average of all the years in the sample.

2.4.3 Artificial Neural Networks

Artificial Neural Networks simulates the human learning process such that the system learns the nature of the relationship between inputs and outputs by repeatedly sampling input/output information sets. Neural networks are characterized by three architectural features: inputs, weights and hidden units; inputs are the data received by the system and grouped, according to assigned weights for their relative importance to each hidden unit. Each hidden unit computes the weighted sum of all inputs and transmits the results to other hidden units which, are performing, simultaneously, similar tasks with their inputs, with interactions continuing until all information is incorporated. The difficulty in using this technique is that it can easily grow to a prohibitive dimension. Lack of transparency is also other major disadvantage of Artificial Neural Networks, so, despite being a usefully tool for prediction, it does nothing to clarify the process and relative importance of the variables.

2.5 Comparing Models

Several comparative studies between different models were made trying to identify which one better estimates the likelihood of default. Shumway (2001) argues that reduced form hazard models are more appropriate than single-period models - static credit scoring models - for forecasting bankruptcy and finds that half the accounting ratios are not statistically significant. Other than that, market size, past-stock returns and idiosyncratic returns variability are all strongly related to bankruptcy. Shumway (2001) proposes a model that uses both accounting ratios and market-driven variables to produce out-of-sample forecasts that are more accurate than those of alternative models. Shumway (2001) details several advantages of hazard models, including the fact that: reduced form hazard models solve the problems of static models by explicitly accounting for time. In a reduced form hazard model, a firm's risk for bankruptcy changes through time and its health is a function of its latest financial data and its age. The bankruptcy probability that a static model assigns to a firm does not vary with time. In econometric terms (1) static models fail to control for each firm's period at risk while hazard models adjust for it automatically; when sampling periods are long, it is important to control for the fact that some firms file for bankruptcy after many years of being at risk while other firms fail in their first year; (2) hazard models incorporate time-varying covariates, or explanatory variables that change with time; (3) hazard models may produce more efficient out-of-sample forecasts by utilizing much more data. The hazard model can be thought of as a binary logit model that includes each firm year as a separate observation. Since firms in the sample have an average of 10 years of financial data, approximately 10 times more data is available to estimate the hazard model than is available to estimate corresponding static models. This data results in more precise parameter estimates and superior forecasts. Sample data was collected from the intersection of the Compustat Industrial File and the CRSP Daily Stock Return File for New York Stock Exchange (NYSE) and American Stock Exchange (AMEX) from 1962 to 1992. A test of models was performed, one testing bankruptcies with market-driven variables exclusively while other combines market-driven variables with two accounting ratios. Combining accounting and market variables results in the most accurate model. This model classifies three-quarters of bankrupt firms in the highest bankruptcy decile, and it only classifies 3.5% of bankrupt firms below the bankruptcy probability median. The model based solely on market-driven variables performs quite well also, classifying 69% of bankrupt firms in the highest probability decile and 95% of bankrupt firms above the probability median. Bankruptcy forecasts can be improved dramatically by conditioning on market-driven variables.

Hillegeist, S. A., D. P. Cram, E. K. Keating, and K. G. Lundstedt (2002) access if two popular accounting-based measures - credit scoring models: Z-score and O-score - effectively summarize public available information about the probability of default. To achieve this goal they developed a measure based on the Black-Scholes-Merton option pricing theory, denoted by BSM-PB. Their test shows that the **structural model outperforms the two credit scoring measures in the amount of relevant information about bankruptcy**.

2.6 Europe's case, particularly Southern Europe Countries

Studies mentioned until now focused, mainly, on samples from US public companies. However, there are several others, where the analysis is made over different locations and type of companies. Several ones focused the analysis of non-listed companies, which, is relevant for the Portuguese case since, these are the large majority in our country - even considering large ones -. This is the case of Christian E. Castro (2008) where he applied a parametric proportional hazard model, in its discrete version, in order to predict probabilities of financial distress (PFDs) on a large dataset of non-listed private Spanish firms between 1994 and 2005. He included four financial dimensions in the analysis of the firm, jointly with controls by sector and size. In addition, he worked over two common factors, in order to study the possible effect of fluctuations in the macroeconomic environment. Surprisingly, it was found that an increase in the real GDP growth rate can generate an increase in the frequency of defaults and estimated probabilities of default (PD). Nevertheless, this effect depends on the age of the firm, being especially important in the case of "young" firms - honeymoon effect -. Some interesting effects have been found, for instance, the effect of GDP growth and interest rates on the estimation of probabilities of default depends of the maturity of the firm. Periods of growth boost the emergence of young firms, with interest rates playing a marginal role in this field. Nevertheless, many of these firms were unable to survive and suffered financial distress after a short period. More mature firms are more exposed to interest rate fluctuations than to GDP growth - this study treated a long growth period in Spanish economy, it´s conclusions could be contradicted in periods of economic recession -.

In Bottazzi G., Grazi I., Secchi A. And Tamagni F. (2009), it is presented an empirical analysis of firm default, exploiting information on distress events occurring in a large panel of Italian firms. The intention was to assess whether the inclusion of economic variables alongside traditional financial ones improves the knowledge regarding the causes of firms' default, possibly trying to increase the chance to correctly distinguish "healthy" firms from those at a risk of distress. The analysis is not limited to publicly traded ones but also to a large group of limited liability firms. Eventually, the interplay between firm idiosyncrasies and market environment has different consequences in profit levels and, ultimately, into exit, default or growth events. Bootstrap probit regressions reveal that economic conditions exert a determinant effect on the probability of default, complementary and additional, with respect to the contribution of financial indicators, and so, findings confirm that adding economic indicators can enhance the understanding of the process leading to firms' default. The final sample for the study included 19.628 Italian manufacturing firms. Accounting data was obtained from CeBI (Centrale di Bilanci), which contains financial statements and balance sheets of virtually all Italian limited liability firms. This data was then matched with a dummy variable 1, for default - at end of 2003 or 2004 -, and 0 otherwise. Default events were provided by the Italian bank only for those firms belonging to their costumer database. Two models were tested; the first one includes, among the regressors', only financial indicators - Interest Expenses over Total Sales, Leverage and Financial-Debt-to-Sales ratio -. The second one, economic variables were added: Size - in terms of total sales -, Labor Productivity - Value Added per employee -, Profitability - return on sales - and Growth - log-difference of Total Sales -.

Robustness checks were performed, namely using Distance to Default (DD). Since this measure derives directly from Merton C. Robert (1974), (DD) applies to publicly traded companies only, and, due to the nature of some of the inputs needed for the model, a solution is to adopt a *naïve* DD measure, as in Bharath T. S. and Shumway T. (2004). Yet, this *naïve* Distance to Default (DD) still requires market data, and so, that was not possible in the context of this work, where only accounting data was available. Nevertheless, the authors built an equivalent measure: bookDD, based on accounting data on the value of shares and debt, which can de derived from available figures on leverage and total assets. The analysis validated the robustness of the findings and, the set of financial and economic characteristics remains valid. The work shows that financial and economic dimensions capture different, albeit complementary, determinants of the process leading to distress. Within financial variables, the cost of debt exerts the most important effect, but economic characteristics also play a role, which is significant over the entire horizon covered by the data analyzed.

In Bonfim D. (2009) a study on the determinants of corporate credit default, taking simultaneously into account firm-specific data as well as macroeconomic information, is developed, over a sample of more than 30.000 Portuguese firms for the period comprised between 1996 and 2002. The results seem to confirm the hypothesis that in periods of economic growth, credit increases and, there may be some tendency to excessive risk-taking. Also, results suggest that idiosyncratic firm characteristics influence default probabilities. Further, firms default history should be taken into account in the assessment given that firms which recorded loan defaults in the recent past seem to display much higher default probabilities than other firms. Finally, when macroeconomic variables are considered the results seem to improve considerably.

Following Basel II implementation and discussion, it was clear that credit risk varies over time and, specifically, it varies with overall macroeconomic conditions. The underlying is that most risk is built up during economic growth periods, when banks apply looser credit standards, and, is materialized only when economy hits a downturn - for detailed discussion see Borio, C.**,** Furfin, C. and Lowe, P. (2001) -. In Rosh, D. (2003) empirical evidence is presented indicating that in economic downturns correlation between borrowers increase as several are affected, simultaneously, by negative systematic shocks.

2.7 Studies on Exit Causes

The academic literature focuses not exclusively in corporate bankruptcies and default but also on different exit forms; being bankruptcies one extreme form of restructuring, firms can find different exit routes; they may be acquired, they may be voluntarily liquidated or they may be merged. Although exit is commonly associated with failure, this may not be the case, as the owners of exiting firms may have made a profit, whether or not business profits have been produced and, furthermore, in liquidations debtors are paid in full. In Bhattacharjee, A., Higson, C., Holly, S., Kattuman, P. (2009) it is made an attempt to identify the factors that increase the likelihood of exit of firms, using data on listed UK companies, from 1965 to 1998. It was estimated a competing-risk model to consider explicitly the joint determination of the probability of being acquired and of being bankrupt. This study confirmed conclusions from prior ones that, exit rates decline with firm age - theoretical models of learning -. Over the business cycle, exit rates increase during downturn. Regarding the impact of macroeconomic instability on exits, there are notable differences in the way in which recently listed firms respond to changes in the macro-environment; there is higher propensity of firms that have been listed during the upturn of the business cycle to go bankrupt as soon as the economy turns down. Firms that overcome the downturn period are more likely to be acquired immediately after the economy enters an up phase. Smooth macroeconomic management is highlighted on the study since sharp variations on inflation and exchange rates affect freshly listed firms adversely.

Mata, J., Antunes, A., Portugal, P., 2011, study the impact of financial variables on bankruptcy and voluntary exit. A combination of two unique datasets was performed. One of the datasets includes information on all firms employing paid labor in Portugal, while another one records all credit relationships between financial intermediaries and non-financial firms. This allows identifying firms that cease to operate separating between those who exit with losses to creditors and those who do not. The restriction of lending relationships to banks, excluding bonds, is not seen as a disadvantage since Portugal – like most European countries – is a bank based economy, in which stock and bond market have a relatively limited scope. The study distinguishes between bankruptcies and voluntary exits analyzing the effect that credit decisions exert upon these two different modes of exit. In a perfect financial markets world firms would be able to borrow the cash they need and only efficient firms would survive in a competitive environment. However, efficient firms may also exit because the owners may lack the funds to keep them going. According to corporate finance theory debt plays important roles: Information asymmetries are alleviated with debt; conflicting interests between shareholders and debt holders; agency theory prediction – debt increases bankruptcy and decreases voluntary exit. Results support the idea that efficiency - measured by size and productivity - is a key driver of the survival of the firms. Larger firms and those that are more productive have lower probabilities of exiting voluntarily and of going bankrupt. Additional findings: firms are cash constrained due to information asymmetries; highly leveraged firms have more probabilities of going bankrupt but are less likely to exit voluntarily; intensive reliance on short-term debt is associated with increased bankruptcy; borrowing from more banks lowers the likelihood of voluntary exit, if the number of banks is small; younger firms exhibit lower chances of exiting voluntary; firms that have more owners and those that have foreign owners have significantly lower probabilities of going bankrupt, while the contrary is true for voluntary exit.

3. STATEMENT OF RESEARCH, METHODOLOGY, CONSTRAINTS

3.1 Research Project

As mentioned initially, **my purpose with this dissertation is to apply bankruptcy and (or) default models to a sample of Portuguese firms that, ideally, would mirror the Portuguese economy in recent year period in which, country is going through a severe economic recession.** In order to pursue that goal, some issues need to be covered: (1) recent financial statements with detailed data, for a group of years, for all firms in the sample; (2) information on bankruptcies and time of occurrence; (3) information on macroeconomic trends for the years of the study; finally, (4) all economic sectors should be present. Achieving these goals requires the availability of the data in order to allow for statistical and econometric treatment. Portuguese corporate tax rules require that all firms submit year-end full financial statements. However, tax authorities and even Statistics Portugal (INE), do not compile data for econometric treatment. Nevertheless, some entities do compile detailed financial information for different purposes namely trade credit risk, trade credit scoring and credit insurance. One of these entities is COFACE® (Portugal). From their web:

"The Coface Group is a trade risk expert and a worldwide leader in credit insurance. It assists companies −regardless of their size, business sector or country− as they grow within their domestic and export countries".

"Coface draws on its worldwide network of credit information entities and CreditAlliance partners. Through this dense web of international information sources, data are fed into an unique risk database, which forms the backbone of coface credit risk rating, management, insurance, and financing offering.
The various types of information obtained by Coface, is cross-checked with data from several hundred public and private sources, and is used to manage each company's score and Coface's risk exposure on a continuous basis." http://www.coface.com/

In order to properly perform their credit insurance and trade credit rating services in Portugal, COFACE®, constantly monitors the corporate sector, collecting all relevant financial information and also other qualitative and financial data. Following their policy of cooperating with academic and university research work, COFACE®, agreed to provide me with a sample from their extensive database, complied for econometric treatment.

3.2 Methodology and Data Description

3.2.1 Method for collecting data

The sample COFACE® agreed to provide needed to abide to some constraints, namely: (1) restricted to 5.000 companies; (2) confidentiality of data – no VAT number or name or any data that would allow the real identification of any entity could be provided; (3) years for financial statements should be limited, between 5 to 6, in order to extract an workable database. To comply with the required restrictions, I asked a sample of the full detail of yearly financial statements, for selected companies, with the following criteria: (1) last reported turnover over or equal to 6.500.000 Euros; (2) last reported number of employees over or equal to 20; (3) exclude bank and insurance companies; (4) years for full detailed financial statement, 2006 to 2011. These selection criteria's resulted in a, approximately, 4.700 firm´s sample – full details will be given ahead.

The data was provided in a Microsoft Access ® format. In order to properly apply econometric models to the sample, I would need to utilize compatible software to import and treat data. The choice was over STATA ®, using as reference manual Cameron A. Colin and Triverdi P. K. (2009).

A note regarding the reliability and quality of the supplied accounting data: In Portugal, firms which surpass 2 of the following figures: (1) balance sheet total: 1.500.000 Euros; (2) Revenues: 3.000.000 Euros; (3) No. employees: 50, are obliged, by corporate governance law, to have a statutory auditor in their board – Revisor Oficial Contas; additionally, all companies in which equity is represented by shares, hold the same obligation, independently of the mentioned limits. From here follows that all firms in our dataset are statutory audited.

3.2.2 Data description

3.2.2.1 Raw Data

As previously mentioned, the sample supplied by COFACE® contained detailed financial statements, from 2006 to 2009, according to Portuguese GAAP – General Accepted Accounting Practices. During this period Portugal accounting GAAP changed, in order to approach international GAAP disclosed in the IFRS/IAS rules, issued by International Accounting Standards Board. This change became compulsory from 2010 onwards. In order for comparability of information, I proceeded with several accounting reclassifications in the financial statements. These accounting reclassifications will be essential for computing the financial ratios that will be part of the final dataset.

Additionally, following fields and information's were requested and included:

FIELD Original Name	Detail and Description
ID	Entity identification for this Sample
CAE	Code for economic activity, according to Portuguese general classification of economic activities
CAE_DESC	Description of CAE
DATA_CONSTITUICAO	Date when company was created
IND_ACCOES_CIVEIS	Dummy for legal court actions against the company for defaulting on payments – usually suppliers. 1 if yes, 0 otherwise.
IND_PROCESSOS_INSOLV	Dummy for company Bankrupt, 1 if yes, 0 otherwise.
IND_DISSOLUCOES	Dummy for company exit. 1 if Yes, 0 otherwise.
ULT_ANOBAL	Last year available financial statements; if company bankrupt, I considered as last year in business.
ULT_ANOBAL_VOLNEG	Last year available financial statements turnover
ULT_ANOBAL_ANEM	Last year financial statements headcount

3.2.2.2 Ratio transformation, additional economic data and final setup

The financial statement is a filtered representation of information in accordance with local GAAP. Many times, companies have incentives to bias the information disclosed in financial reporting for different purposes: income tax, leverage, covenant restriction, profit and equity enhancement, etc. The use of financial ratios as a tool for evaluating past company performance and economic position has a long tradition in academic literature and corporate performance analysis; this includes corporate failure prediction – see Altman I. Eduard (1968). As an easy-to-use tool, it enables analysts to gain insights over current financial position and forecast future results. Nevertheless, important limitations need to be taken in consideration, particularly: (1) benchmark values may not be directly compared over different industries; (2) the use of different accounting methods; (3) different ratios may provide contradictory results over same company. For further details see CFA® Program Curriculum (2011).

Many studies on bankruptcy prediction try to identify ratios that were expected to capture relevant aspects for the analysis; besides the work already mentioned by Altman I. Eduard (1968) also Beaver, W. (1966) tested several ratios, in several bankruptcy cases, identifying "working capital funds flow/total assets" and "net income/total assets" as best discriminators. Ohlson (1980) was the first to apply logit analysis to bankruptcy prediction, concluding that reporting power appears to be inferior than reported in previous studies. For a survey on different approaches to bankruptcy prediction see Morris (1988); for an example on using ratios for bankruptcy prediction see Bernhardsen E. (2001). So, a fundamental initial step is to extract from the raw database the information and variables that will be tested for relevance as bankruptcy predictors. Based on prior papers, particularly, the already mentioned: Bottazzi G., Grazi I., Secchi A. And Tamagni F. (2009); Castro, E. Christian (2008); Bonfim, D. (2009); the option fell over the following set of ratios (all ratios calculated with final year book values); missing values, namely for years after bankruptcy or exit, appear in STATA® as ".".

Profitability Ratios:

- *ROA (Return On Assets):* (NETPROFITy/ASSETSy)*100; ASSETS as Balance Sheet Total; depending on Net Profit ROA can be positive or negative;
- *ROE (Return On Equity):* (NETPROFITy/EQUITYy)*100; can be both positive or negative. Both Equity and Net Profit can take negative values, interchangeably, which may lead to misleading interpretations. Hence, proper adjustments are needed to avoid these biases. In the case where both numerator and denominator are negative we would obtain, without any adjustment, a positive percentage indicating a profitable performance for the period, obviously, an absurd result. The chosen approach was to turn the ratio negative whenever Net Profit is negative and positive otherwise. In detail: (1) if Net Profit is negative and Equity is also negative, multiply result

by -1; (2) if Net Profit positive and Equity negative, multiply by -1; (3) in all remaining cases, do nothing.

- *ROS (Return On Sales)*: (NETPROFITy/SALESy)*100; SALES amount considers the net sum of sales of products with revenues from charged services, in accordance with Portuguese GAAP; depending on Net Profit ROS can be positive or negative;
- *ErnPower (Earnings Power):* (EBITy/ASSETSy)*100; can be positive or negative depending on the nature of EBIT – Earnings Before Interest and Taxes, the result considering only business operation;
- *BusinessMargin:* (EBITy/SALESy)*100; Can be positive or negative depending on EBIT nature.
- *AssetsTurnover:* SALESy/ASSETSy; it is a productivity and efficiency measure; can only take positive values;
- *Interest Coverage:* (EBITy/INTERESTEXPENDITUREy); INTEREST EXPENDITURE the total expenditure on interest over bank loans and overdrafts, bonds and shareholder loans. Depending on the nature of EBIT can be both positive or negative;

Solvency and Liquidity Ratios:

- *SOLVENCY:* (EQUITYy/ASSETSy)*100; Can be both positive or negative depending on Equity;
- *FINANCIAL LEVERAGE:* ASSETSy/FinancialDEBTy; Financial Debt, the sum of bank loans, bonds and shareholders loans; short and long term debt is considered; the idea is to identify the several possible sources that a firm can use to finance itself, particularly, interest paying debt; excludes working capital. Can only take positive values;
- *LIQUIDITY:* CurrentASSETSy/CurrentLIABILITIESy; Current Assets considers Cash, Receivables, Inventories and prepaid expenses; Current Liabilities considers Accounts payable, short-term debt and long term debt due in one year, Accruals, Leases until one year maturity. Can only take positive values;

In addition to the mentioned ratios, I chose to add additional controls for test:

- *YearsActivity*: Firm´s age;
- *Maturity*: dummy taking 1 if firm “immature” as defined by age <= 10 years; 0 otherwise;
- *Dimension*: log(ASSETS);

The age and maturity of firms is found to be a significant control for distress prediction. Thornhill S. and Amit R. (2003) found systematic differences between young and mature firms in determining failure. In López-Garcia P. and Puente S. (2006) similar results are disclosed using a sample of Spanish firms. Dimension of firm is also found of relevance in several studies, see Bhattacharjee, A., Higson, C., Holly, S., Kattuman, P. (2009); Eklund, T., Larsen, K. and Bernhardsen, E. (2001) and Bunn, P. and Redwood, V. (2003).

Finally, conclusions on works that served as starting point for this thesis, confirm that, prediction results are substantially improved – ex. Bonfim, D. (2009) - with the introduction of systematic factors, that influence simultaneously all firms, alongside with idiosyncratic firm characteristics. In continuum, it was choose to include the following economic variables (Source: IMF's World Economic Outlook):

- *GDPVAR*: Annual percentages of constant price GDP are year-on-year changes;
- *Investment*: Expressed as a ratio of total investment, in current local currency, and GDP in current local currency. Total Investment or gross capital formation is measured by the total value of the gross fixed capital formation and changes in inventories and acquisitions less disposals of valuables for a unit or sector;
- *Inflation*: Annual percentages of average consumer prices are year-on-year changes.

For a description of the sample data in STATA® final setup, see appendix A.

3.2.2.3 Summary Statistics

The dimension of the microeconomic dataset used in this dissertation work is of 4.714 firms; as I will be detailing ahead, not all firms will present observations for all years, as some will disappear due to bankruptcy and(or) exit.

For further detail on the nature of firms in the dataset, see bellow table 1:

Master CAE	CAE Description	No. Firms	No. Employees	Bankruptcy			Exit		Civil/Court Actions	
				No. Firms	Bankruptcy Frequency %	No. Employees	No. Firms	No. Employees	No. Firms	Frequency %
A	Agricultura, produção animal, caça, floresta e pesca	44	5.286	4	9,1%	574	0	0	24	54,5%
B	Indústrias extractivas	22	3.233	1	4,5%	58	0	0	15	68,2%
C	Indústrias transformadoras	1.445	251.136	67	4,6%	8.921	29	4.288	741	51,3%
D	Electricidade, gás, vapor, água quente e fria e ar frio	23	7.433	0	0,0%	0	0	0	18	78,3%
E	Captação, tratamento e distribuição de água; saneamento, gestão de resíduos e despoluição	100	18.003	2	2,0%	213	5	1.036	71	71,0%
F	***Construção***	***394***	***82.922***	***109***	***27,7%***	***17.169***	***17***	***2.840***	**349**	**88,6%**
G	Comércio por grosso e a retalho; reparação de veículos automóveis e motociclos	1.600	221.129	76	4,8%	10.255	42	4.675	933	58,3%
H	Transportes e armazenagem	288	84.380	6	2,1%	937	2	654	231	80,2%
I	Alojamento, restauração e similares	95	40.476	4	4,2%	1.152	0	0	63	66,3%
J	Actividades de informação e de comunicação	150	39.633	4	2,7%	409	10	1.771	94	62,7%
K	Actividades financeiras e de seguros	35	1.894	0	0,0%	0	0	0	14	40,0%
L	***Actividades imobiliárias***	***28***	***2.428***	***4***	***14,3%***	***265***	***1***	***117***	**21**	**75,0%**
M	Actividades de consultoria, científicas, técnicas e	137	26.886	8	5,8%	772	3	577	78	56,9%
N	Actividades administrativas e dos serviços de apoio	201	181.376	15	7,5%	12.275	7	2.680	126	62,7%
O	Administração Pública e Defesa; Segurança Social	5	2.048	0	0,0%	0	1	108	2	40,0%
P	Educação	12	1.744	0	0,0%	0	0	0	6	50,0%
Q	Actividades de saúde humana e apoio social	103	97.327	1	1,0%	150	5	13.670	70	68,0%
R	Actividades artísticas, de espectáculos, desportivas e recreativas	25	4.939	3	12,0%	359	0	0	18	72,0%
S	Outras actividades de serviços	7	1.174	0	0,0%	0	0	0	5	71,4%
	Total	**4.714**	**1.073.447**	**304**	**6,45%**	**53.509**	**122**	**32.416**	**2.879**	**61,07%**

Table 1 - Dataset Summary

In the table 1, information is presented regarding nº. of firms, employees, bankruptcies, exits and court actions, grouped by economic activity, according to the Portuguese general classification of economic activities. Some highlights are immediately possible: (1) construction sector is the one suffering the most severe turmoil, losing 27, 7% of firms to bankruptcy. (2) real-estate activities follow construction on the percentage of bankruptcies; the two activities are closely linked, and are strongly subject to systematic factors that affect economy. With the Portuguese economy severely hit by recession since 2009 - since then only one year presented a slight growth: 2010 - construction and real-estate sectors are experiencing a severe contraction since crisis started.

Chava S. and J. A. Jarrow (2004) show that most of the academic literature regarding bankruptcy prediction models is based on datasets containing at most 300 bankruptcies; at this point our dataset is slightly superior, presenting 304 events. Looking proportionally, we are working with a distress rate of 6,45% over the total number of firms, similar to other studies, see Duffie D., Saita L. and Wang K. (2007). The dataset also includes information regarding exit; again, construction has the higher relative number of exits, which represent 2, 59% of total. Our dataset reveals 4.288 living healthy firms, with are 90, 9% from total.

In Table 2 generic firm characteristics in the dataset are displayed:

In Euros (ex. age & No. Employees)	Mean	Max	Min	Median	Std Deviation
Assets	58.400.568	18.755.807.464	0	11.074.060	458.368.814
Net Profit	1.763.224	5.889.053.625	-397.198.623	193.653	40.846.831
Equity	18.229.613	10.592.493.823	-2.759.276.232	3.208.101	184.338.937
Sales	40.146.393	9.574.706.375	85	13.407.952	193.848.530
Age	26	255	1	21	21
No. Employees	228	22.734	10	86	733

Table 2 - Firms Generic Statistics

From the procedure used to collect the dataset, we may infer, with some safety, that this sample represents the subset of the largest Portuguese companies.

In the mean column we may have a picture of the typical large company, considering the dimension of the Portuguese economy and of the corporate sector. A brief comment on Assets, particularly the zero amount in Min column; the fact that the dataset includes firms that went bankrupt or exited business their information, for the following years after the event, appears as zero in all fields related to financial statements information.

In table 3 we find the main economic indicators. We confirm the recession period Portugal faces since 2008, with the exclusion of 2010, where a slight growth on GDP was achieved:

Years values in %	GDP Anual Growth	Investement	Inflation
2006	1,44	23,134	3,043
2007	2,386	22,828	2,423
2008	-0,008	23,152	2,651
2009	-2,507	19,916	-0,903
2010	1,331	18,998	1,391
2011	-2,159	17,587	3,443

Table 3 - Generic Economic Indicators

Table 4 details some statistics for selected ratios:

Ratio	mean	max	min	median	sd	iqr	kurtosis	skewness
ROA	2,9	33	-32	2,1	8,3	5,7	7,3	-0,34
ROE	0,55	99	-490	7,3	60	17	32	-4,7
ROS	1,9	38	-55	1,5	10	4,3	15	-1,7
ErnPower	4,4	38	-25	3,1	8,7	7,3	6,9	0,66
BusinessMargin	3,8	43	-51	3,2	11	6,1	12	-0,96
AssetsTurnover	1,7	9,3	0,052	1,3	1,5	1,2	12	2,6
InterestCoverage	954	58.500	-647	3	6.492	12	69	8,1
Liquidity	1,8	10	0,00023	1,4	1,5	0,96	16	3,2
Solvency	32	100	-43	30	23	28	3,8	0,082
FinancialLeverage	86	4.611	0,042	4	515	6,4	67	7,9

note: iqr: interquartile range = p75 - p25

Table 4 - Financial Ratios Statistics

Some notes need to be mentioned regarding the construction of the above ratios: (1) in all ratios, a set of outliers – in both extremes – were cancelled, using the 99th and 1st percentile – when appropriate, by replacing observations above the 99th percentile with its value and repeating similar procedure for observations bellow 1st percentile; the above statistics were computed after the cancelling of the outliers; (2) missing values were left as blank – ex. Years after bankruptcy, since, STATA® can exclude this values from any calculation.

The table is disclosing important information from this dataset that, as we see, can be considered as a strong representation of the larger firms in Portugal, and, probably, the ones more solid and prepared to face market competition. In this sense, it may not be excessive, to extend the conclusions from this study, as representing Portugal enterprise sector. To provide support for this inference we see that the companies in this dataset **represent about 22% of the total employed population** in Portugal, considering also public services (Source: Statistics Portugal - www.ine.pt; last available data for total employed population, 2010 – 4.948.800) . All ratios are found to be leptokurtic, with SOLVENCY being the exception, revealing strong peak around the mean and fat tails. Low levels of profitability, particular ROE: 0,55%. Additionally, besides strongly leptokurtic, observations on ROE are negatively skewed, meaning that the majority of them fall below the mean. Similar results for ROA and ROS, but here, higher profitability and lower negative skew; these contradictory results on the mean amounts result from the higher dispersion in ROE observations. Results on SOLVENCY, with 32% average, reveal the known tendency for reliance on leverage from Portuguese firms. A Kurtosis near 3 and skewness near 0, reveal observations very closed of being normally distributed.

For a graphical representation of the data I use the Kernel density plot. When data takes many observations, a better alternative to a histogram is a Kernel density plot, see appendix B. The Kernel is formed with the weighted values calculated with the kernel function K:

$$\widehat{f}_K = \frac{1}{qh}\sum_{i=1}^{n} w_i K\left(\frac{x - X_i}{h}\right),$$

That places greater weight on points x closer to X_i. The kernel function utilized for appendix B, was the Epanechnikov:

$$K[z] = \begin{cases} \frac{3}{4}(1 - \frac{1}{5}z^2)/\sqrt{5} & \text{if } |z| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases}$$

3.2.3 Constraints

Due to the type of data collected for the dataset, the main restriction is the impossibility of applying structural and reduced form models; these, use information contained in publicly traded securities – shares, bonds, which is not present in this dataset. This limits the possibilities of studies to the use of credit scoring models.

4. ECONOMETRIC METHODOLOGY

4.1 Test on Means

	Mean Values [BankRuptcy (Y/N)]						
	0=NO	1=YES	t-ratio	Degrees of Freedom	diff = mean(0) - mean(1)	Ha: Diff not 0; Pr(|T| > |t|)	Means are Significantly Different?
ROA	3,114	-1,009	20,932	1.817,880	4,123	0,000	YES
ROE	1,937	-21,581	10,748	1.660,650	23,518	0,000	YES
ROS	2,111	-2,222	14,624	1.719,630	4,333	0,000	YES
Liquidity	1,824	1,450	12,459	1.975,890	0,374	0,000	YES
Solvency	33,166	20,719	25,367	1.892,890	12,447	0,000	YES
Earnings Power	4,566	1,358	16,939	1.875,760	3,208	0,000	YES
Business Margin	4,019	0,818	10,580	1.730,870	3,201	0,000	YES
Assets Turnover	1,740	1,240	17,100	1.947,940	0,501	0,000	YES
Interest Coverage	1.004	188	8,947	2.541,710	816,156	0,000	YES
Financial Leverage	90,591	17,675	12,992	4.377,390	72,916	0,000	YES
Age	26	27	-1,081	2.195,370	-0,465	0,280	NO
Dimension -log(Assets)	16,347	16,597	-8,424	1.890,120	-0,250	0,000	YES

Table 5 - Tests on Means

Being the main purpose of this thesis trying to understand what leads to corporate bankruptcy, one first step would be to test mean differences between bankrupt firms and the remaining ones. Table 5 lists mean values for the two groups. Logic would lead to conclude that financial and economic situation between the two groups of firms would differ, and, indeed, differences are substantial, with **bankrupt group presenting lower profitability ratios – even negative means for ROA, ROE and ROS – and being more dependent on financial debt – lower Solvency and Financial Leverage means.** Similarities appear only in Liquidity and Assets Turnover ratios. Contradicting conclusions on other studies, age and dimension are similar between these two groups.

For testing more accurately if indeed these variables are different between the two groups, table 5 presents results on mean differences tests that use the following test statistic:

$$t = \frac{\overline{x} - \overline{y}}{\left(s_x^2/n_x + s_y^2/n_y\right)^{1/2}}$$

,

This tests statistic applies when variances are unknown and different between two groups of firms, which, is the expected in this case. The result is distributed as a t-student with the following degrees of freedom (Welch formula):

$$-2+\frac{\left(s_x^2/n_x+s_y^2/n_y\right)^2}{\frac{\left(s_x^2/n_x\right)^2}{n_x+1}+\frac{\left(s_y^2/n_y\right)^2}{n_y+1}}$$

All tests confirm, with the exception of age, that means between the two groups are indeed statistically different. For details on STATA® commands see appendix C.

4.2 Regression Analysis with Binary Outcome Models

Being the purpose of the dissertation the understanding of corporate bankruptcy prediction, which is, by nature, a qualitative problem, these types of problems can be treated by regression analysis in applied statistics. Models for mutually exclusive binary outcomes or dichotomous variables focus on the determinants of the probability *p* of the occurrence of one outcome rather than the alternative outcome that occurs with a probability of 1 – *p.*

4.2.1 Correlation and Other Relevant Issues

Prior to moving forward to regression analysis, we should analyze a correlation matrix between variables. Particular attention should be given to **multicollinearity,** that happens when **two or more predictors in a regression are highly correlated**. When two variables have an exact linear relationship between the two, they are called to possess **Collinearity**. Multicollinearity refers to a situation in which two or more explanatory variables in a multiple regression model are highly linearly related. In the presence of multicollinearity, the estimate of one variable's impact on the dependent variable while controlling for the others tends to be less precise than if predictors were uncorrelated with one another. Multicollinearity does not reduce the predictive power or reliability of the model as a whole, at least within the sample data themselves; it only affects calculations regarding individual predictors. Up to some degree, collinear variables contain the same information about the dependent variable, and, if nominally "different" variables contain the same information, or are highly correlated then, they are redundant. The best regression models are those in which the predictor variables each correlate highly with the dependent (outcome) variable but correlate at most only minimally with each other. For additional details see Gujarati, Damodar. (2003).

CORRELATION Matrix	IND_PROCESSOS_INSOLV (BankRuptcy)	IND_ACCOES_CIVEIS (Court Actions)	ROA	ROE	ROS	Liquidity	Solvency	ErnPower	Business Margin	Assets Turnover	Interest Coverage	Financial Leverage	Maturity	Dimension	GDPVAR	Investment	Inflation
IND_PROCESSOS_INSOLV (BankRuptcy)	1,00																
IND_ACCOES_CIVEIS (Court Actions)	0,17	1,00															
ROA	-0,12	-0,12	1,00														
ROE	-0,10	-0,07	0,63	1,00													
ROS	-0,10	-0,07	0,71	0,49	1,00												
Liquidity	-0,05	-0,08	0,24	0,11	0,21	1,00											
Solvency	-0,13	-0,15	0,49	0,29	0,40	0,47	1,00										
ErnPower	-0,09	-0,10	0,81	0,49	0,53	0,16	0,35	1,00									
Business Margin	-0,07	-0,04	0,64	0,42	0,77	0,16	0,31	0,60	1,00								
Assets Turnover	-0,09	-0,10	0,09	0,06	-0,01	-0,08	-0,11	0,11	-0,09	1,00							
Interest Coverage	-0,02	-0,03	0,16	0,06	0,10	0,10	0,10	0,14	0,11	0,01	1,00						
Financial Leverage	-0,04	-0,02	0,09	0,04	0,07	0,13	0,11	0,06	0,04	0,03	0,13	1,00					
Maturity	-0,04	-0,04	-0,06	-0,05	-0,09	-0,10	-0,21	-0,04	-0,07	0,11	0,00	0,01	1,00				
Dimension	0,05	0,21	-0,02	-0,03	0,06	0,04	0,07	-0,05	0,11	-0,50	0,01	0,02	-0,11	1,00			
GDPVAR	0,02	0,01	0,07	0,07	0,04	-0,02	-0,01	0,20	0,03	0,02	0,01	0,01	-0,04	-0,05	1,00		
Investment	0,04	0,02	0,06	0,06	0,03	0,04	-0,01	0,11	0,00	0,02	0,00	0,03	-0,04	-0,07	0,61	1,00	
Inflation	-0,02	-0,01	0,00	0,00	0,00	-0,05	0,01	0,20	0,01	0,02	-0,01	-0,01	-0,03	-0,02	0,40	0,22	1,00

Table 6 - Correlation Matrix

Table 6 depicts the correlation matrix between variables in the dataset. Correlation coefficient is obtained trough:

$$\rho_{X,Y} = \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

Results are limited between -1, in the case of a perfect decreasing (negative) linear relationship (anticorrelation), and 1, in the case of a perfect positive (increasing) linear relationship (correlation).

High positive correlation coefficients are found between some profitability variables, in detail (I am considering coefficients above 0,50 as high): ROE – ROA: 0,63; ROS – ROA: 0,71; Earnings Power – ROA: 0,81; Business Margin – ROA: 0,64; Earnings Power – ROS: 0,53; Business Margin – ROS: 0,77; Business Margin – Earnings Power: 0,60. For the remaining set of variables, the only additional high correlation coefficient is found between Investment – GDPVar: 0,61. The results are not surprising, since, profitability ratios are extracted using lines from the same Profit & Loss account, and so, strongly interconnected; similar justification may be given for the result in Investment with GDP variation, since, investment level is one GDP component.

An important result regards the correlation with, what would be our future dependent variable for bankruptcy analysis: IND_PROCESSOS_INSOLV; **none of the remaining variables are highly correlated with bankruptcy dummy variable, being all, except for one, negatively correlated.** As mentioned above, this may jeopardize the quality of regression models using this set of variables, and observations, present in the dataset obtained for this dissertation thesis, as in other similar studies where accounting ratios cannot perfectly predict bankruptcy.

4.2.2 Binary Outcome Models

There are several binary outcome models which share a common structure. The purpose of these models is to fit a set of explanatory variables to a binary or dichotomous dependent variable, answering questions like, yes or no type or, a group of mutually exclusive ones – polychothomous. There are three approaches to developing a probability model for a binary response variable - see Gujarati, Damodar. (2003), chapter 15, for additional details:

- The Linear Probability Model;
- The Logit Model;
- The Probit Model

4.2.2.1 The Model

The purpose would be to fit a linear function of explanatory variables: x_{i1} , x_{i2} , …, x_{in} , and the random term μ , such that:

$$y_i^* = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \mu_i$$

$$y_i = 1 \; if \; y_i^* > 0$$

$$y_i = 0 \; else$$

The dependent outcome variable, *y,* has a Bernoulli, or binomial distribution with one tail, and takes one of two values, with a probability of p :

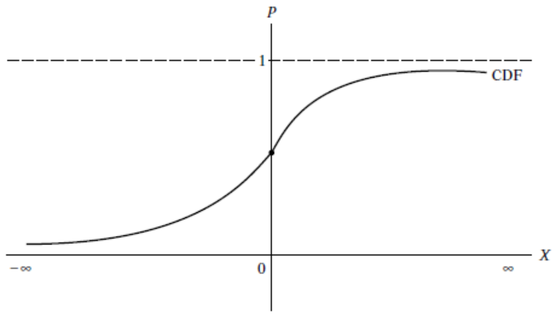
$$y = \begin{cases} 1 \text{ with probability } p \\ 0 \text{ with probability } 1 - p \end{cases}$$

The probability mass function for the observed outcome, y , is $p^y(1-p)^{1-y}$, with $E(y) = p$ and $Var(y) = p(1-p)$.

A regression model is formed by parameterizing *p* to depend on an index function $x'\beta$, where x is a *K* regressor vector and β is a vector of unknown parameters. In binary outcome models, the conditional probability has the form:

$$p_i \equiv \Pr(y_i = 1|x) = F\,(x_i'\beta)$$

where $F\,(\cdot)$ is a specified parametric function of $x'\beta$, usually a cumulative distribution function that ensures that the probability lies between 0 and 1, and that it varies nonlinearly with x on $(-\infty, \infty)$; geometrically the cumulative distribution function (c.d.f.) of a random variable resembles a sigmoid or S-shaped curve:



The practical question now is: which cumulative distribution function (c.d.f.)? For although all (c.d.f.) are S shaped, for each random variable there is a unique (c.d.f.). For historical as well as practical reasons, the (c.d.f.) commonly chosen to represent the 0–1 response are (1) the logistic function and (2) the normal function, the former giving rise to the logit model and the later to the probit model. The linear probability model corresponds to linear regressions and does not impose the $0 \leq p \leq 1$ restriction being so less powerful than logit and probit models.

Binary outcome models can be given a continuous unobservable – or latent – interpretation. We can distinguish between the observed binary outcome, *y*, and an underlying continuous unobservable – or latent – variable, y^* , that satisfies the single-index model:

$$y^* = x'\beta + \mu \quad (1)$$

we observe

$$y_i = 1 \; if \; y_i^* > 0$$
$$y_i = 0 \; else \quad (2)$$

The zero threshold is a normalization of no consequence. Given the latent-variable models (1) and (2), we have

$$\Pr(y = 1) = \Pr(x' + \mu > 0)$$

$$= \Pr\,(-\mu < x'\beta)$$

$$= \mathrm{F}(\mathrm{x}'\beta)$$

where $F(\cdot)$ is the cumulative distribution function (c.d.f.) of $-\mu$.

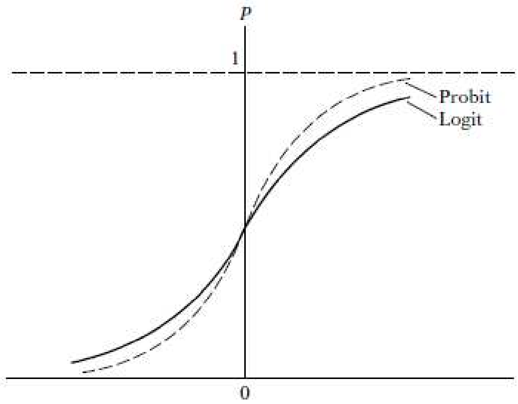
For the Logit and Probit model, estimation is fit by maximum likelihood (ML). For N independent observations, the maximum likelihood estimation (MLE), $\hat{\beta}$, maximizes the associated log-likelihood function:

$$Q(\beta) = \sum_{i=1}^{N} [y_i lnF(x_i'\beta) + (1 - y_i)ln\{1 - F(x_i'\beta)\}]$$

The MLE is obtained by iterative methods and is asymptotically normally distributed. Details from Cameron A. Colin and Triverdi Pravin K. (2009), chapter 14 and Gujarati, Damodar. (2003), chapter 15.

4.2.2.2 Models Comparison

Between Logit and Probit, which model should we choose? In The majority of applications the models are very similar. The main difference regards the fatter tails in the logistic function, that is to say, the conditional probability p_i approaches zero or one at a slower pace in logit than in probit:



Therefore, there is no fundamental reason why to choose logit or probit. Many researchers choose logit for its comparative mathematical simplicity.

4.2.2.3 Panel-Data Particularities

In panel-data the same cross-sectional unit - a firm in our dataset – is surveyed over time, therefore, panel-data have space as well as time dimensions. There are some advantages on using panel-data: (1) they increase sample size considerably; (2) by studying repeated cross-section observations, panel data are better suited to study the dynamics of change; (3) more sophisticated behavioral models are able to be studied trough panel-data. Despite substantial advantages, some issues need to be addressed regarding panel-data: cross-sectional data need to account for heteroscesdasticy and in time series data for autocorrelation. Panel-data methods require additional sophistication, since, the standard errors of panel data estimators need to be adjusted because each time-period of data is not independent from prior ones. STATA® has several procedures – xt – for treating panel data. Appendix E provides details on STATA® panel-data commands and statistics regarding our dataset.

4.2.2.4 Notes on Fixed Effects and Random Effects Models

Considering panel data, the linear function of explanatory variables: x_{i1} , x_{i2} , …, x_{in} , and the random term μ , would be:

$$y_{it} = \alpha_i + \beta_1 x_{i1t} + \beta_2 x_{i2t} + \cdots + \beta_k x_{ikt} + \mu_{it}$$

Being y_{it} the dependent variable with regressors x_{it}, where α_i is an individual effect, i , denotes the individual and t denotes time; terms suggests that the intercepts on cross-sectional data are different but still assume that the slope coefficients are constant across cross-sectional data, in this way taking in account each firm "individuality". This means that, although the intercept may differ across individuals, each individual's intercept does not vary over time, being *time invariant*.

In a **Fixed Effects (FE)** model, the individual-specific effect,α_i, is treated as an unobserved random variable that may be correlated with the regressors. In long panels this poses no problem. Alternatively, the **Random Effects (RE)** model, the individual-specific effect is treated as an unobservable random variable with a normal distribution. So, which of the optional models should be choose, FE or RE? The literature considers that if individual-specific effects are correlated with regressors then FE would be the one to choose. In the case of our dataset, we can expect individual-specific effects – ex. Management Ability – to be correlated with financial ratios regressors, since, management ability is determinant for profitability. Gujarati, Damodar. (2003), adds that the FE model is the appropriate when the cross-section is large and the number of time units is small if we believe that the cross-sectional units in our dataset are not random drawings from a larger sample.

Nevertheless, efficient Fixed Effects (FE) estimators rely on within variation; in our dataset, within variation for our dependent variable is 0, as firms only go bankrupt once, thus limiting the application of FE models to our dataset. Bellow, summary statistics for our dependent variable redrawn from SATA®:

```
. xtsum IND_PROCESSOS_INSOLV
```

Variable		Mean	Std. Dev.	Min	Max	Observations
IND_PR~V	overall	,0644888	,2456259	0	1	N = 28284
	between		,2456476	0	1	n = 4714
	within		0	,0644888	,0644888	T = 6

4.2.2.5 Logit Model Estimation

Following the discussion in 4.2.2.4., we will go forward applying Logit model with Random-Effects to our dataset. STATA ® utilizes the following Random-Effects (RE) Logit model, fitted by maximum likelihood:

$$\Pr(y_{it} \neq 0|x_{it}) = \mathrm{P}(x_{it}\beta + \alpha_i)$$

α_i are i.i.d., $N(0, \sigma_\alpha^2)$, and P is the standard logistic distribution: $P(z) = \{1 + \exp(-z)\}^{-1}$

For a detailed discussion, see Chamberlain, Gary. (1980) and Cameron A. Colin and Triverdi Pravin K. (2009), chapter 18.

4.2.3 Model Results and Comments

4.2.3.1 Results Obtained with Binary Outcome Models

Taking into account prior statistics and tests applied over our dataset, namely: correlation matrix and tests on means – see tables 5 and 6 -, I will start by performing estimations using a limited set of variables that will increase in number for each estimation. The purpose will be the understanding of the value each added variable bring in to the model.

XtLogit (1)	Coeficient	z-ratio	p-value P>|z|
ROA	-0,029813	-2,44	0,015
Solvency	-0,0288414	-5,55	0,000
Maturity	-1,261795	-3,9	0,000
Investment	0,0895979	2,4	0,016
Constant	-10,95721	-13,48	0,000

Table 7 - Logit Model(1)

Table 7 depicts a first set of explanatory variables for estimation. It comprises a profitability indicator: Return on Assets (ROA), a leverage indicator: Solvency, dummy for age: Maturity, and a systematic variable: Investment. Excluding Investment, all variables are exhibiting negative coefficients meaning that: increasing these variables, decreases bankruptcy probability. This is the anticipated and reasonable conclusion regarding ROA and Solvency: companies that are more profitable and less indebted should be more economically solid and less prone to bankruptcy. Relative surprise obtained in Maturity and Investment results. Studies and practice suggest that mature firms have lower probability of default and bankruptcy; Castro, E. Christian (2008) along with other studies identified the maturity of firm has a strong effect on bankruptcy indentifying a “honeymoon” effect, for young firms that are not able to survive.

```
Summary for variables: Maturity
     by categories of: IND_PROCESSOS_INSOLV (Court Action (or) if Banckrupt; =1 if yes)
```

IND_PROCESSOS_INSOLV	mean	sd	max	min
0	,1798186	,3840435	1	0
1	,1217105	,3270408	1	0
Total	,1760713	,3808875	1	0

In the figure above, some summary statistics for Maturity grouped by bankrupt and non-bankrupt companies. Mean and Standard Deviations are very similar between the two groups; in 4.1, we saw that the mean average age of firms in two groups in our dataset is practically equal, apparently contradicting conclusions from other studies. Two reasons may justify this contradiction: (1) due to the nature of our dataset, Portuguese small companies are excluded from the analysis, and the great majority of start-ups fall in this excluded group; (2) the economic turmoil started in 2008 is affecting Portuguese economy as a whole, destroying companies that not long ago were of considerable dimension and age - see construction. Bellow is shown the count of companies “born” after 2005, adding to 412 and from these, 13, “died” (3,2%) trough bankruptcy:

```
Summary for variables: Maturity
     by categories of: IND_PROCESSOS_INSOLV (Court Action (or) if Banckrupt; =1 if yes)
```

IND_PROCESSOS_INSOLV	N
0	399
1	13
Total	412

Results on Investment are also somehow surprising and contradictory, since, taking account of the positive sign of the coefficient, increasing Investment economic indicator would be associated with an increase in corporate bankruptcy probability. From table 3 we may confirm that Investment has been decreasing, year over year, in the period covered by our dataset, influencing, in this way, all observations and firms, not distinguishing between bankrupt and no bankrupt.

Considering p-values, all variables are statistically significant, at a 5% significance level.

XtLogit (2)	Coeficient	z-ratio	p-value P>|z|
ROA	-0,0293058	-2,13	0,033
Solvency	-0,0320195	-5,28	0,000
Financial Leverage	-0,0012507	-2,33	0,020
Maturity	-1,185426	-3,36	0,001
Dimension	0,2666771	3,47	0,001
GDPVAR	-0,0017036	-0,03	0,974
Investment	0,1050658	2,21	0,027
Constant	-15,64695	-9,22	0,000

Table 8 - Logit Model(2)

Adding Financial Leverage, Dimension and GDP variation, we get the set of variables depicted in table 8. Regarding variables from xtLogit(1), no material differences have occurred and described conclusions hold satisfactorily for xtLogit(2) for those variables. For the added variables (1) Financial Leverage presents a negative coefficient, meaning: the higher the ratio, the lower the bankruptcy probability; Financial Leverage for this dataset was calculated trough the ratio ASSETSy/FINANCIAL DEBTy, expressing that, higher the result, the less a firm's indebtedness; we should expect that firm's with increasing dependence on debt, hold an increasing risk for default and are more prone for a bankruptcy event – normally triggered by debt holders and(or) suppliers with accounts receivables; the negative coefficient on Financial Leverage confirms this expectation. (2) Dimension results seem contradictory with common sense; the positive coefficient is indicating the larger the firm – size measured by the book value of assets – the larger the bankruptcy probability. Even several bankruptcy models identified size as a significant factor. In Bernhardsen E. (2001), a study over Norwegian companies, the same variable utilized in this study for Dimension – log(Assets) – was found without statistical significance. (3) GDP year-on-Year variation, has a more common sense result, however is not statistically significantly different from zero, and so, it is not possible to properly interpret its negative coefficient, that would reveal that a period of economic growth, lowers firm's bankruptcy probability. Our dataset, contrary to the Investment variable, shows periods of positive and negative growth.

With the exception of GDPVAR, all added variables are statistically significant at a 5% significance level. GDPVAR with a p-value of 0,925 is very far from significance.

XtLogit (3)	Coeficient	z-ratio	p-value P>|z|
ROA	-0,0306479	-1,28	0,202
ROE	-0,0003471	-0,24	0,811
Liquidity	-0,0291382	-0,31	0,756
Solvency	-0,0310797	-4,88	0,000
ErnPower	-0,0022319	-0,11	0,914
Business Margin	0,0028805	0,28	0,782
Financial Leverage	-0,0012279	-2,37	0,018
Maturity	-1,281075	-3,35	0,001
Dimension	0,2730956	3,47	0,001
GDPVAR	-0,0039203	-0,07	0,942
Investment	0,1086017	2,28	0,023
YearsActivity(AGE)	-0,0049697	-0,8	0,421
Constant	-15,65011	-9,15	0,000

Table 9 - Logit Model(3)

Table 9 shows results on XtLogit(3), where five new variables were added to the model. (1) ROE (Return on Equity): comments are very similar to the ones written for ROA earlier, but with ROA holding a much relevant coefficient; as we seen in 4.2.1, ROA and ROE are highly corrected (0,63), meaning that the information added to the model is somehow redundant, the two variables are, together, insignificant at a 5% level, and, ROA p-value substantially increased from xtlogit(2); (2) Liquidity: ratio that discloses the firms solvability to comply with immediate payments; results are somehow anticipated, namely regarding the negative sign of the coefficient; nevertheless, with a large p-value, it is far from relevant statistic significance; Liquidity also has a strong correlation with Solvability (0,47) which may indicate redundant information added to the model. (3) ErnPower (Earnings Power): highly correlated with ROA (0,84) ; very high p-value makes this variable statistically insignificant. (4) Business Margin: highly correlated with prior profitability variables, adding redundancy to the model, ROA (0,64) and Earnings Power (0,60); similarly to Earnings Power, high p-value for statistical significance; presenting positive coefficient, despite very low and near zero, we should expect a negative coefficient in line with remaining profitability variables. (5) YearsActivity (AGE): comments for age are equivalent to the ones written for maturity; as we saw in 4.1, the mean age between the two groups of companies is very similar; the variable is not statistically significant at a 5% level.

XtLogit (4)	Coeficient	z-ratio	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	9,07094	20,88	0,000
ROA	-0,0273954	-0,84	0,399
ROE	-0,0005988	-0,34	0,732
ROS	-0,0017483	-0,09	0,93
Liquidity	-0,0544616	-0,47	0,64
Solvency	-0,0279828	-3,47	0,001
ErnPower	0,0143843	0,54	0,587
Business Margin	0,0019687	0,11	0,914
Assets Turnover	-0,7570599	-4,35	0,000
Interest Coverage	-9,24E-06	-0,29	0,770
Financial Leverage	-0,0011115	-1,01	0,313
Maturity	-1,227243	-2,77	0,006
Dimension	-0,1347315	-1,19	0,236
GDPVAR	0,0239584	0,34	0,733
Investment	0,0984545	1,69	0,09
Inflation	-0,0890961	-1,16	0,245
YearsActivity(AGE)	-0,001529	-0,19	0,852
Constant	-18,54127	-7,71	0,000

Table 10 - Logit Model(4)

Finally, all variables are introduced in regression in XtLogit(4), table 10. From prior models, results are similar, however, regarding new variables, some highlights need to be mentioned. (1) Comments for profitability variables follow the ones in prior models; new profitability variables are producing similar results as prior ones; ROS – Return On Sales and Interest Coverage as, ROA, ROE, Earnings Power and Business Margin, are not statistically significant; excluding Interest Coverage, all profitability variables are highly correlated - see 4.2.1 for details; the interesting issue and, somehow surprising, on the results of profitability variables, is their lack of statistic significance for bankruptcy; we should expect a clear difference in profitability between the two groups: bankrupt and non-bankrupt. From 4.1 we tested and confirmed this difference, which would forecast a different result for the regression models. In spite of that, we should attend to characteristics of Portuguese corporate sector, in which several firms survive, for long and repeated periods, with negative results. It follows that, some firms with negative results, within some or all years in the dataset, should be present; I display a detailed statistic, printed from STATA®, considering all observations with negative Net Profit:

```
. tabstat ROS if NetProfit < 0, statistics( mean count ) by(IND_PROCESSOS_INSOLV)

Summary for variables: ROS
     by categories of: IND_PROCESSOS_INSOLV (Court Action (or) if Banckrupt; =1 if yes)

IND_PROCESSOS_INSOLV |      mean         N
---------------------+--------------------
                   0 | -9,658542      4587
                   1 | -12,43253       495
---------------------+--------------------
               Total | -9,928736      5082
------------------------------------------
```

From all observations with negative profits, 90,3% fall in the non-bankrupt group. Surely, this proportion, is affecting the Binary Regression Models conclusions regarding profitability variables; the economic downturn period also justifies this proportion. (2) **Assets Turnover** is a productivity ratio and resulted statistically significant for the model; also has one of the **most relevant coefficient** and with an expected negative sign, meaning: **higher productivity and efficiency, lower bankruptcy probability**. (3) Inflation has a negative coefficient, indicating periods of increased inflation lower bankruptcy probabilities; Inflation may indicate economic growth, resulting in increased turnover from companies; in our dataset the average annual inflation was 2%; Inflation is not statistically significant at a 5% and 10% levels. (4) IND_ACCOES_CIVEIS, this dummy is the most relevant variable in the model, the highest positive coefficient (9,07), low p-value (0,00) indicating statistic relevance, showing that the existence of **legal actions against a firm is of strong relevance for bankruptcy prediction**; this result is anticipated since, when a firm starts delaying – or defaulting on payment terms, ex. to suppliers, employees, on loans – one of the first decisions that firms management take is to move forward with legal actions against the defaulting firm. (5) Investment loses statistic significance at a 5% level; remains significantly different from zero at a 10% level of significance.

Appendix F details the STATA® output for prior models.

4.2.3.2 Comparing Results

Confidance level:	95%							
Dependent Var.:	IND_PROCESSOS_INSOLV	Wald Test H0: β(k) = 0						
Model	Log. Likelihood	Test Stat: Chi2	p-value:Prob >= chibar2	Rho	Test Of Rho	chibar2(01)	Nº Observ.	Nº Groups
XtLogit (1)	-1.483,52	74,96	0,00	0,93	0,00	8.305,38	26.657	4.714
XtLogit (2)	-1.412,76	80,78	0,00	0,94	0,00	7.580,65	22.599	4.495
XtLogit (3)	-1.409,33	81,33	0,00	0,94	0,00	7.509,17	22.366	4.485
XtLogit (4)	-1.161,10	486,46	0,00	0,97	0,00	6.827,82	21.407	4.425

Table 11 - Tests on Logit Models

Table 11 presents some hypothesis tests so we may confirm and compare the several fitted models. Wald test for the null hypothesis that the coefficients of all regressors are equal to zero, meaning, there is no interaction effects between variables; for all models the null hypothesis is rejected – *p-values = 0.*

4.2.3.3 Control for Economic Downturn Period

During the period analyzed on this dissertation, Portugal faced a severe economic downturn period, triggered by an international financial turmoil. We saw in figure 4 that this downturn started in started in 2008 – the first years with negative GDP growth – after two years of slight economic growth. In this sense, this dataset allows a division of the analysis between a two year economic growth period, 2006 and 2007, averaging 1,913% annual GDP growth, and, a crisis period, from 2008 to 2011, averaging -0,84% annual GDP growth. This possibility of division is a main feature of this dissertation comparing to similar studies. One first step is to extract from the dataset, the quantity of firms that went bankrupt before and after the crisis period; for this we need to know in which year the company left business through bankruptcy. In the dataset we have information on the last year firms publicized financial statements, this being an annual tax obligation for all firms in Portugal, and, not fulfilling will bring them penalties. Firms in Portugal comply with this obligation annually and besides detailed financial statements, complete information regarding other taxes – ex. VAT – and labor headcount is disclosed. This obligation stops when a firm is dissolved or bankrupt. As we mentioned when describing the data, COFACE extract this data from all Portuguese companies for their database, from where the present dataset was extracted. This information is publicly available in: www.portaldaempresa.pt .

In the dataset, the field: ULT_ANOBAL indicates the last year of available financial statements, and, linking with the dummy variable for bankruptcy, we may consider as being the year company went bankrupt.

Firms Bankrupt	Count
2006	0
2007	0
2008	0
2009	62
2010	95
2011	147
Total	304

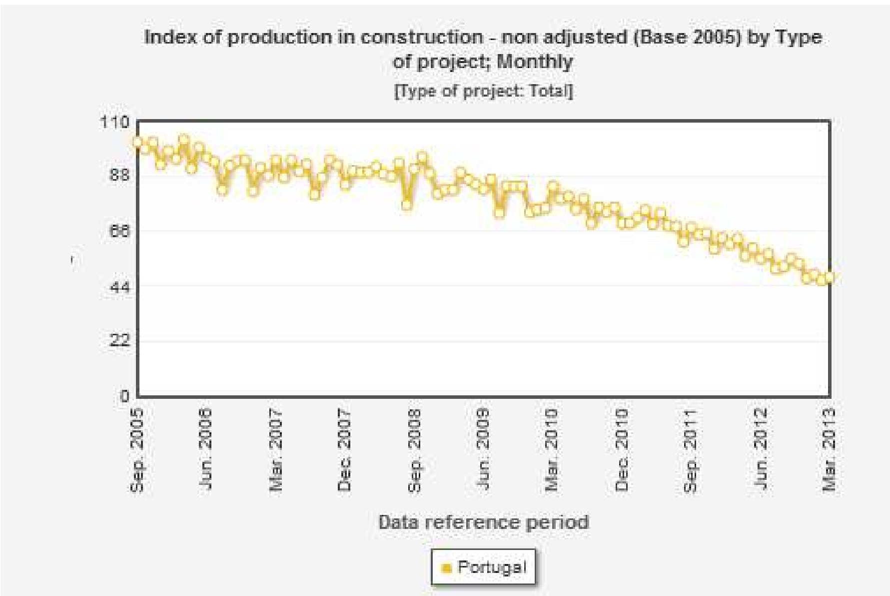
Table 12 - Bankruptcies Count by Year

Table 12 details the count of firms that went bankrupt according to their last year of available financial statements. The first and immediate conclusion is that in our dataset there are no firms bankrupt before 2009, or before the start of the crisis period. This peculiar result may be consequence of the process used to select firms for the dataset, detailed in 3.2.1.; due to the constraints imposed on the selection, and the option on choosing the larger companies in Portugal, bringing information from each firm for all the years in the study, criteria for selection was dimension, and, not state of business operation; in doing so, for large companies in 2006 that are not bankrupt at the time, that state will appear in later period.

This limits our analysis since the next logical step would be a statistical test over means for companies bankrupt before and after crisis started. Even so, a dummy variable was created: Crisis, taking 0 for years before 2006 and 2007 and 1 otherwise. Appendix F details a XtLogit model that includes the Crisis dummy variable. The Crisis dummy reveals to be not statistically significant.

4.2.3.4 Results for Construction Sector

As we saw in 3.2.2.3 construction was one corporate sector that suffered severely with bankruptcies events, 35,9% of the 304 bankruptcy events in our dataset came from construction sector, the one with more bankruptcy events. Portuguese economy was, still is, highly dependent on construction activity, being decisive for GDP and employment figures in Portugal. The sector is being affected by a severe business contraction particularly during the period analyzed in our data set.



The above graph shows the evolution of construction activity in Portugal since 2005 (Source: INE – Statistics Portugal www.ine.pt). We see that, presently, activity level is less than 50% of 2005 figures, and the fall is being accentuated from end of 2009 onwards.

In Table 13 a presentation of a logit model considering only bankruptcies in construction sector – the MASTER_CAE =”F”:

XtLogit (Construction)	Coeficient	z-ratio	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	14,33807	8,49	0,000
ROA	-0,0885807	-0,6	0,547
ROE	-0,0208092	-1,83	0,067
ROS	0,1310495	1,63	0,102
Liquidity	-1,063433	-2,81	0,005
Solvency	-0,0677808	-2,29	0,022
ErnPower	0,0991067	0,82	0,414
Business Margin	-0,063804	-1	0,319
Assets Turnover	-0,8678032	-0,89	0,374
Interest Coverage	-5,13E-05	-1,28	0,202
Financial Leverage	-0,0011389	-1,13	0,260
Maturity	1,531219	1,06	0,288
Dimension	0,9663313	3,04	0,002
GDPVAR	0,1688003	0,6	0,547
Investment	0,2242721	0,94	0,345
Inflation	-0,3471554	-1,2	0,231
YearsActivity(AGE)	0,0343485	2,66	0,008
Constant	-35,84991	-4,41	0,000

Table 13 - Logit Model(Construction)

Conclusions are similar to the models that considered the full dataset – XtLogit(4) –however some important differences need to be emphasized: (1) Assets Turnover and Investment loose statistical significance holding economic interpretation; (2) Liquidity and Solvency gain importance in the model, being liquidity, now, the second most important coefficient in the model, revealing that increasing liquidity, lowers bankruptcy probability, and achieves statistically significance; Solvency gains economic relevance increasing its negative coefficient from XtLogit(4); (3) Age – YearsActivity and Dimension are now statistically relevant however, with a positive sign, indicating that older and larger companies have increased chances of going bankrupt; this somehow contradictory result may be explained by the nature of our dataset, that includes only large companies and the contraction that is affecting construction in Portugal is affecting all players including mature and large ones. Appendix F includes STATA® results for this model.

4.2.3.5 Lagged Model

In studies for determinants of default or bankruptcy, past performance is evaluated in order to understand how it can affect default probability, which may be useful in predicting future defaults; it can be relevant to assess if a firm will be in stress in the near future considering its present situation, see Bonfim, D. (2009).

XtLogit (L1)	Coeficient	z-ratio	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	21,36102	21,36	0,000
ROA	-0,0223205	-0,48	0,628
ROE	0,0000874	0,03	0,978
ROS	-0,0032968	-0,14	0,891
Liquidity	-0,1458588	-0,92	0,356
Solvency	-0,0392761	-3,26	0,001
ErnPower	0,0196637	0,51	0,607
Business Margin	0,0055289	0,24	0,810
Assets Turnover	-1,081952	-4,42	0,000
Interest Coverage	-1,12E-05	-0,21	0,837
Financial Leverage	-0,0010605	-1,32	0,188
Maturity	-1,817883	-2,84	0,005
Dimension	-0,3156724	-2,06	0,039
GDPVAR	0,0254391	0,16	0,87
Investment	0,1107979	0,74	0,458
Inflation	-0,1084676	-0,4	0,686
YearsActivity(AGE)	-0,0140273	-1,33	0,184
Constant	-33,15359	-7,83	-41,450

Table 14 Logit Model(lagged)

In table 14 is depicted a model similar to XtLogit(4) but now variables were lagged by one year. Results obtained are similar to the ones in XtLogit(4), with the following differences: (1) Dimension p-value lowers and is now statistically significant at 5% level (2) Investment p-value increases significantly and variables looses statistic significance even at a 10% level; (3) Age lowers p-value however still insufficient for gaining statistical significance. Appendix G shows STATA® results.

4.2.3.6 Firm's Exit Analysis

Firms may leave business trough different processes. Bankruptcy can have important consequences namely for the ones who bear the costs; voluntary exit or liquidations, which may be associated with failure, need not to be as so, since the owners may have made profits. In liquidations, independently of the existence of losses, debt holders are fully paid, contrasting with bankruptcies where, at least in some cases, debtors may be left unpaid. In Bhattacharjee, A., Higson, C., Holly, S., Kattuman, P. (2009), a study over the determinants of business failure and exit in large UK firms, the purpose was to examine the influence of the macroeconomic cycle in business exit and failure. They found different behaviors according to firms maturity, particularly young firms that have been created during economic upturn periods that are more likely to go bankrupt when economic downturn arrives; companies that survive downturns are more likely to exit, through mergers, when in an upturn economic period. Mata, J., Antunes, A., Portugal, P. (2011) study bankruptcy and voluntary liquidation patterns in Portuguese firms; results support mainly two conclusions (1) efficiency is a key driver for firm's survival; larger and more productive firms are less likely to exit or go bankrupt. (2) Cash constrains due to information asymmetries strongly influence bankruptcies and exits.

Our dataset has information regarding voluntary exit, in the dummy variable IND_DISSOLUCOES, with 1 for YES and 0 for NO.

	Mean Values [Exit (Y/N)]						
	0=NO	1=YES	t-ratio	Degrees of Freedom	diff = mean(0) - mean(1)	Ha: Diff not 0; Pr(|T| > |t|)	Means are Significantly Different?
ROA	2,870	2,889	-0,054	596,984	-0,019	0,957	NO
Solvency	32,515	28,643	4,336	601,785	3,872	0,000	YES
Assets Turnover	1,710	1,730	-0,317	593,588	-0,019	0,751	NO
Financial Leverage	84,468	145,632	-1,958	478,562	-61,164	0,051	YES
Age	27	21	9,372	788,838	6,069	0,000	YES
Dimension -log(Assets)	16,359	16,440	-1,499	601,255	-0,081	0,134	YES

Table 15 - Tests on Means in Exit Group

Table 15 details some means and tests on means on some financial dimensions from the dataset. Differences that should be highlighted appear in Solvency and particularly in Financial Leverage; Results on Financial Leverage are somehow surprising, since exit companies appear with a much stronger ratio than non-exit ones; one reason may be that these companies are deleveraging, repaying debt, immediately before being dissolved. Contrary to bankruptcy, profitability and productivity ratios are quite similar between both groups; however, in ROA and Assets Turnover we do not reject the null hypothesis of means being equal. Appendix H details STATA® screens for means and test on means differences.

The same logic of analysis needs to be applied as for bankruptcy study and a binary outcome model can be applied having Exit dummy variable as the outcome variable. Table 16 details results from the application of the Logit model:

XtLogit (Exit)	Coeficient	z-ratio	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	-0,6753947	-1,34	0,181
ROA	0,0385765	0,67	0,504
ROE	-0,0006676	-0,16	0,869
ROS	0,0107779	0,29	0,774
Liquidity	0,0384529	0,22	0,828
Solvency	-0,028468	-2,03	0,043
ErnPower	-0,0185971	-0,43	0,667
Business Margin	-0,0140295	-0,45	0,653
Assets Turnover	-0,2471506	-1,16	0,246
Interest Coverage	8,48E-06	0,24	0,811
Financial Leverage	0,0001497	0,46	0,645
Dimension	-0,0222252	-0,11	0,911
GDPVAR	0,0108031	0,07	0,941
Investment	0,1437542	1,16	0,248
Inflation	-0,1212642	-0,72	0,472
YearsActivity(AGE)	-0,0374279	-2,31	0,021
Constant	-29,77124	-6,81	0,000

Table 16 - Logit Model(exit)

Results on this dataset are significantly different from the ones obtained in bankruptcy analysis. (1) There is no variable with a significantly influencing coefficient, all are bellow 1; (2) There are only two statistically significant coefficients at 5%: Solvency and Age; (3) Wald Chi2 = 15,95, with a p-value of 0,4565, meaning: we do not reject the null hypothesis of all coefficients being zero at a 5% and 10% significance level. Contradicting studies mentioned before, controls for dimension and productivity are not found to be statistically significant, even at a 10% significant level.

4.2.3.7 Analyzing Results on Default

The variable IND_ACCOES_CIVEIS is a dummy for the existence of court actions against a firm. These situations arise due to different kinds of conflicts where a firm is being charged for some flaw or default. The start of a legal action in order to manage a conflict is always an ultimate form of resolution, usually after a previous negotiation ending without any agreement. The great majority of these types of charges are related to defaults on scheduled payments, commonly to suppliers, banks, employees, etc. The substance of these legal conflicts is the existence of some kind of default on some type of obligation or compromise.

From table 1 we see that 61% of the firms in our dataset are subjected to, at least, one of these charges. We may test this dummy as dependent variable, that is to say, we may analyze if some of our set of variables may, in some way, be driving defaults. The procedure will be similar to the one applied in bankruptcy study, first analyzing and testing difference in mean between the two groups and second to apply a Logit model.

	Mean Values [Legal Charges (Y/N)]						
	0=NO	1=YES	t-ratio	Degrees of Freedom	diff = mean(0) - mean(1)	Ha: Diff not 0; Pr(|T| > |t|)	Means are Significantly Different?
ROA	4,216	2,008	21,277	21.845,600	2,208	0,000	YES
ROE	6,659	-3,365	14,052	25.202,100	10,024	0,000	YES
ROS	2,922	1,173	13,745	23.319,300	1,749	0,000	YES
Solvency	37,226	29,359	27,032	21.028,400	7,868	0,000	YES
Business Margin	4,528	3,384	8,499	22.888,500	1,144	0,000	YES
Assets Turnover	1,885	1,599	14,751	18.978,000	0,286	0,000	YES
Financial Leverage	103,481	75,430	3,781	15.179,800	28,051	0,000	YES
Age	26	27	26,716	22.427,700	-0,593	0,021	YES
Dimension -log(Assets)	15,990	16,599	-36,666	24.108,200	-0,609	0,000	YES

Table 17 - Mean Differences (Legal Charges)

Above table 17 shows means differences between firms that hold legal charges against them and others who don't. Additionally, table 17 provides the results for the tests regarding the statistical significance for the same mean differences. We see that firms subjected to legal charges against have lower profitability and are more leveraged that the opposite group. Ratios ROA, ROE Solvency and Financial Leverage are all significantly higher in the no Legal Charges group and all are statistically significant at 5% significance level. These results point that healthier enterprises are better shielded against legal conflicts. Slight difference on Assets Turnover but still favouring the No Legal Charges group, and, also, difference is statistically significant. Results on Age and Dimension are similar to the ones tested when studying bankruptcy. Details on STATA® results may be found in Appendix J.

Fitting a Logit model similar to the one used for bankruptcy and exit analysis we obtain the results presented on table 18. The most relevant variables are Maturity and Dimension, being both statistically significant. The most intriguing result comes from Dimension that has the most influential coefficient, however, being positive is indicating that, the larger the enterprise, the larger the probability of being subjected to legal charges. This apparently counter-intuitive conclusion may be related to: (1) the larger the company, the larger the number of stakeholders with which company interacts, increasing the probabilities for conflicts to arise; (2) The method used for selecting companies for this dissertation thesis, that is to say, biased towards larger Portuguese companies. None of the remaining variables present similar influence as the two previously mentioned and, excluding Solvency and Investment, all are not statistically significant. Details on STATA® results may be found in Appendix J.

XtLogit (Default)	Coeficient	z-ratio	p-value P>|z|
ROA	-0,0534342	-1,45	0,147
ROE	-0,0037263	-1,27	0,206
ROS	-0,0191631	-0,95	0,34
Liquidity	-0,0977601	-0,8	0,426
Solvency	-0,0727171	-7,35	0,000
ErnPower	-0,0006075	-0,02	0,985
Business Margin	0,0279256	1,5	0,134
Assets Turnover	0,2100608	1,68	0,092
Interest Coverage	-2,46E-05	-1	0,315
Financial Leverage	0,0002315	1	0,317
Maturity	-1,615108	-2,92	0,003
Dimension	2,046212	15,72	0
GDPVAR	0,0008666	0,01	0,996
Investment	0,2648569	2,26	0,024
Inflation	-0,0580259	-0,4	0,686
YearsActivity(AGE)	-0,0175626	-1,41	0,158
Constant	-33,63375	-9,01	0,000

Table 18 - Logit Model (default)

4.2.4 Conclusions on Binary Models

With the purpose of trying to understand what drives bankruptcy, some sophisticated tests and regression models were applied to the present dataset. Table 17 lists the results from the most relevant model. The dataset includes idiosyncratic variables as systematic ones, affecting all firms in the market. Confirming prior studies, some variables were found relevant: Solvency, productivity, measured by Assets Turnover. Being able to increase these ratios seems to act as shield against bankruptcy. Surprisingly, none of the profitability variables proved statistical significance. From the systematic variables that were tested only Investment resulted in statistic relevance; from all, the most important coefficient was from IND_ACCOES_CIVEIS indicating, as is perfectly anticipated, this as a strong predictor for bankruptcy. The same model was tested using only companies in construction sector due to the relevance this sector has in Portugal; here the main difference was the importance gained by Liquidity and Solvency as shields against bankruptcy. A similar model was applied using as outcome variable Exit, but not so relevant conclusions were obtained as for bankruptcy. Appendix I depicts STATA® results. Due to the information available on the dataset no conclusions could be extracted regarding the influence of the crises period on the data.

	XtLogit (4)			XtLogit (Construction)			XtLogit (Lagged)			XtLogit (Exit)		
XtLogit Models	Coeficient	z-ratio	p-value P>|z|	Coeficient	z-ratio	p-value P>|z|	Coeficient	z-ratio	p-value P>|z|	Coeficient	z-ratio	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	9,07094	20,88	0,000	14,33807	8,49	0,000	21,36102	21,36	0,000	-0,6753947	-1,34	0,181
ROA	-0,0273954	-0,84	0,399	-0,0885807	-0,6	0,547	-0,0223205	-0,48	0,628	0,0385765	0,67	0,504
ROE	-0,0005988	-0,34	0,732	-0,0208092	-1,83	0,067	0,0000874	0,03	0,978	-0,0006676	-0,16	0,869
ROS	-0,0017483	-0,09	0,93	0,1310495	1,63	0,102	-0,0032968	-0,14	0,891	0,0107779	0,29	0,774
Liquidity	-0,0544616	-0,47	0,64	-1,063433	-2,81	0,005	-0,1458588	-0,92	0,356	0,0384529	0,22	0,828
Solvency	-0,0279828	-3,47	0,001	-0,0677808	-2,29	0,022	-0,0392761	-3,26	0,001	-0,028468	-2,03	0,043
ErnPower	0,0143843	0,54	0,587	0,0991067	0,82	0,414	0,0196637	0,51	0,607	-0,0185971	-0,43	0,667
Business Margin	0,0019687	0,11	0,914	-0,063804	-1	0,319	0,0055289	0,24	0,810	-0,0140295	-0,45	0,653
Assets Turnover	-0,7570599	-4,35	0,000	-0,8678032	-0,89	0,374	-1,081952	-4,42	-1,562	-0,2471506	-1,16	0,246
Interest Coverage	-9,24E-06	-0,29	0,770	-5,13E-05	-1,28	0,202	-1,12E-05	-0,21	0,837	8,48E-06	0,24	0,811
Financial Leverage	-0,0011115	-1,01	0,313	-0,0011389	-1,13	0,260	-0,0010605	-1,32	0,188	0,0001497	0,46	0,645
Dimension	-0,1347315	-1,19	0,236	0,9663313	3,04	0,002	-0,3156724	-2,06	0,039	-0,0222252	-0,11	0,911
GDPVAR	0,0239584	0,34	0,733	0,1688003	0,6	0,547	0,0254391	0,16	0,87	0,0108031	0,07	0,941
Investment	0,0984545	1,69	0,09	0,2242721	0,94	0,345	0,1107979	0,74	0,458	0,1437542	1,16	0,248
Inflation	-0,0890961	-1,16	0,245	-0,3471554	-1,2	0,231	-0,1084676	-0,4	0,686	-0,1212642	-0,72	0,472
YearsActivity(AGE)	-0,001529	-0,19	0,852	0,0343485	2,66	0,008	-0,0140273	-1,33	0,184	-0,0374279	-2,31	0,021
Constant	-18,54127	-7,71	0,000	-35,84991	-4,41	0,000	-33,15359	-7,83	0,000	-29,77124	-6,81	0,000

Table 19 - Logit Models Compare

Finally I fitted a Logit model for testing default. The proxy used for default was the binary variable for the existence of legal charges against each firm. Starting by analysing, testing and confirming that means on several rations are indeed different between the two groups of companies, the results on Logit regression reveal maturity and Dimension as the most relevant variables in the model.

4.3 Survival Analysis

The purpose of survival analysis is trying to understand the time to the occurrence of an event of interest - see Cleves M., Gould W., Gutierrez R., Marchenko Y. (2008) for details and study on Survival Analysis -, in our case: Bankruptcy. We saw previously in Shumway (2001) that reduced form hazard models are more appropriate than static credit-scoring models in the way they explicitly account for time. In a reduced form hazard model, a firm's risk for bankruptcy changes through time and its health is a function of its latest financial data and its age.

4.3.1 Survival and Hazard model

4.3.1.1 Hazard and Survival Function

If we call *T* nonnegative random variable indicating time to a failure event. Survival analysis refers to *Ts* as survivor function, *S(t)*, or its hazard function *h(t)*. The survivor function is the reverse of the cumulative distribution function of *T* and reports the probability of surviving beyond time *t:*

$$S(t) = 1 - F(t) = \Pr\,(T > t)$$

S(t) is the probability that there is no failure until *t.* The function is equal to one ate *t = 0,* and decreases towards zero as *t* approaches infinite. The hazard function, *h(t),* is the instantaneous rate of failure, and, being a rate, it has units 1/*t.* It is the probability that the failure event occurs within an interval, conditional upon the subject having survived to the beginning of that interval, divided by the width of the interval:

$$h(t) = \lim_{\Delta t \to \infty} \frac{\Pr\,(t + \Delta t > T > t|T > t)}{\Delta t} = \frac{f(t)}{S(t)}$$

It can vary from zero, meaning no risk at all, to infinity, meaning the certainty of failure at that instant. It is the underlying process that determines the shape of the hazard function. If we say a hazard rate is 2/day, then, we are saying, were the rate to continue for an entire day, we would expect two failures. Hazard rates, were they to stay constant, can have an alternative interpretation: hazard rates have units *1/t;* hence, the reciprocal of the hazard has units *t* and represents how long we would expect to have to wait for a failure if the hazard rate stayed constant. If the hazard rate is 2/day, then, we would expect to wait half a day for a failure.

4.3.1.2 Hazard Model

Survival analysis is concerned with the time to the occurrence of an event. If we fit a linear regression trough Ordinary Least Squares (OLS) method, the difficulty would be the assumed normality distribution of the residuals ε_i ; this assumed normality of time to an event is unreasonable for many events. At her core, survival analysis main focus is to substitute the normality assumption with something more appropriate for the problem in hand. Three approaches are used to deal with this problem: (1) parametric survival analysis where a more reasonable distribution assumption for ε_i is applied; (2) Semi-parametric modeling, where assumptions on the distribution of failure times are not required, since events occur at given times. These can be ordered and the analysis can be performed exclusively on the ordering of survival time being this the parametric component of the analysis. The effect of the covariates is still assumed to take a certain form; (3) Non-parametric analysis would abandon all assumptions and follow a philosophy of “letting the data speak for itself”.

A hazard parametric model is usually written as:

$$h_j(t) = g(t, \alpha + x_j\beta_k)$$

That is, the hazard – intensity in which an event occurs – for subject *j* is some function *g()* of $\alpha + x_j\beta_k$ where we allow the presence of multiple predictors through the row vector x_j in which case β_k is a column vector of regression coefficients. This notation can also incorporate the semi-parametric model:

$$h_j(t) = h_0(t)\exp\,(\alpha + x_j\beta_k)$$

Where $h_0(t)$ is called the baseline hazard, that is, the hazard subject j faces is the same hazard everyone faces modified by x_j. Particularly, the above model is called the proportional hazard model, since subject j faces multiplicative proportional hazard and function exp() is chosen to avoid the problem of $h_j(t)$ becoming negative.

4.3.2 Survival-Time Data

4.3.2.1 Data Censoring

In real data analysis situations, we often do not know when failures occurred, at least for every subject present in the dataset. Censoring is defined as when a failure event occurs and the subject is not under observation. We can think of censoring a something caused by a censor standing between us and reality that prevent us of seeing the exact time of the event we know occurs. During the study period, subjects are enrolled and data are collected for a follow-up period, or, the period under which the subject is under observation. Data collection stops when a subject fails, the study ends or the subject leaves the study for other reasons. However the subject can come at risk of failure before or afterwards the enrollment. So, censoring can be of different types: (1) Right Censoring when the subject participates in the study for a period and, thereafter, is no longer observed; (2) Interval Censoring when we do not know the exact failure time, all we know is that failure occurs between two known time points; (3) Left Censoring when failure event occurred before subject was under observation.

4.3.2.2 Survival-Time and Data in STATA®

Before applying hazard models to our dataset we need to declare in STATA® survival-time data. Particularly, in our case, data needs to be organized as survival-time or time-span – other option: count-data. In survival-time data, the observations represent periods and contain three variables that record the start time of the period, the end time, and an indicator of whether failure or right-censoring occurred at the end of the period. The representation of the response of these three variables makes survival data unique in terms of implementing the statistical methods in the software.

The command used to declare survival-time in STATA® is *stset.* In *stset* we declare the (1) failure event; (2) Origin: when a subject becomes at risk; (3) Enter: specifies when a subject first comes under observation, in our dataset the variable EnterYear was generated being equal to StartYear when higher than 2006 and equal 2006 to all remaining situations; (4) Exit: specifies the latest time the subject is both under risk and observation. Appendix L details STATA® results on stset applied on our dataset. Appendix L also includes some STATA® generic statistics for survival – data in our dataset.

4.3.3 Fitting Regression Models

In STATA® the command *stcox* fits, via maximum likelihood, proportional hazard models on survival-time data. *stcox* fits the Cox proportional hazard model, where the hazard is assumed to be:

$$h(t) = h_0(t)\exp\left(\beta_1 x_1 + \cdots + \beta_k x_k\right)$$

This model was first presented by Cox, D. R. (1972). Stcox obtains parameters estimates, $\hat{\beta}$, by maximizing the partial log-likelihood function:

$$Log\ L = \sum_{j=1}^{D}\left[\sum_{i\epsilon D_j} x_i\beta - d_j log\left\{\sum_{k\epsilon R_j}\exp\left(x_k\beta\right)\right\}\right]$$

Being x_i the row vector of covariates for the time interval $(t_{01}, t_i]$ for the *ith* observation in the dataset, *j* ,indexes the ordered failure times $t_{(j)}$, D_j is the set of d_j observations that fail at $t_{(j)}$; d_j is the number of failures at $t_{(j)}$; and R_j is the set of observations *k* that are at risk at time $t_{(j)}$.

4.3.3.1 Cox Regression Results

An identical approach will be used as the one applied for binary outcome analysis, starting with simpler models and gradually increasing complexity.

Table 18 shows the results of fitting a firs Cox model using only controls regarding firms' idiosyncrasies regarding profitability, leverage and productivity:

	Hazard Ratios			Coefficients		
StCox(1)	ratio	z-stat.	p-value P>|z|	Coeff.	z-stat.	p-value P>|z|
ROA	0,9758105	-6,36	0	-0,02449	-6,36	0
Solvency	0,9782295	-14,04	0	-0,02201	-14,04	0
Assets Turnover	0,6780594	-12,3	0	-0,38852	-12,3	0

Table 20 - Cox Model(1)

Two sets of results are displayed in StCox(1): one regarding hazard ratios and, other for coefficients. Lower hazard ratios imply longer survival times. Assets Turnover is revealing to be the most important control for increasing survival time, having the lowest hazard ratio and coefficient, making the strongest contribution to lower final hazard result.

In table 19 StCox(2) results are depicted; additional idiosyncratic controls were added, nevertheless the general conclusion is similar to the one described for StCox(1). Particularly, the stronger influence for increasing survival-time is productivity measured by Assets Turnover. Excluding ROE, all ratios are statistically significant at a 5% level; contradicting results on Logit binary models, here, profitability ratios are statistically significant.

	Hazard Ratios			Coefficients		
StCox(2)	**ratio**	**z-stat.**	**p-value P>|z|**	**Coeff.**	**z-stat.**	**p-value P>|z|**
ROA	0,985307	-2,98	0,003	-0,0148	-2,98	0,003
ROE	0,999314	-1,78	0,075	-0,00069	-1,78	0,075
Solvency	0,977549	-13,13	0	-0,02271	-13,13	0
Assets Turnover	0,636775	-11,09	0	-0,45134	-11,09	0
Financial Leverage	0,999556	-2,59	0,009	-0,00044	-2,59	0,009
Dimension	0,939959	-2,31	0,021	-0,06192	-2,31	0,021
YearsActivity(AGE)	0,926666	-3,59	0	-0,07616	-3,59	0

Table 21 - Cox Model(2)

Finally a Cox model was fitted considering all firms idiosyncratic controls in the dataset; table 19 details results for StCox(3):

	Hazard Ratios			Coefficients		
StCox(3)	ratio	z-stat.	p-value P>|z|	Coeff.	z-stat.	p-value P>|z|
IND_ACCOES_CIVEIS (Court Actions)	7,037094	15,86	0,000	1,9512	15,86	0,000
ROA	0,97892	-2,61	0,009	-0,02131	-2,61	0,009
ROE	0,999231	-1,94	0,053	-0,00077	-1,94	0,053
ROS	1,005069	1,05	0,292	0,00506	1,05	0,292
Liquidity	1,003521	0,13	0,9	0,00351	0,13	0,9
Solvency	0,98108	-10,15	0	-0,0191	-10,15	0
ErnPower	1,004155	0,6	0,548	0,00415	0,6	0,548
Business Margin	1,000214	0,05	0,961	0,00021	0,05	0,961
Assets Turnover	0,625649	-10,7	0	-0,46897	-10,7	0
Interest Coverage	1,00E+00	-0,09	0,925	-1,09E-06	-0,09	0,925
Financial Leverage	0,999186	-2,37	0,018	-0,00081	-2,37	0,018
Dimension	0,865831	-5,07	0	-0,14407	-5,07	0
YearsActivity(AGE)	0,927414	-3,48	0	-0,07535	-3,48	0

Table 22 - Cox Model(3)

Conclusions do not differ substantially from to the ones in Logit models, particularly regarding the new ratios added in StCox(3): Liquidity, ROS, Earnings Power, Business Margin and Interest Coverage are all not statistically significant. Also, hazard ratios are higher than one which may be somehow contradictory for corporate financial analysis since, increasing this variables, would lead to higher hazard lowering survival time; additionally, these ratios are multiplied with positive coefficients; however, contrary to Logit models, ROA and ROE are statistically significant. Similar as in the Logit regression model, there is strong relevance regarding the presence of court actions against the firm; this dummy variable is an important control for increasing hazard and lowering survival time in the model applied to this dataset. Age and dimension, contrary to the Logit regression results, are now statistically significant, being relevant as protection against hazard failure. Appendix M details STATA® results on all mentioned Cox models.

Following IND_ACCOES_CIVEIS the main control is Assets Turnover, confirming the relevance efficiency has in managing a company. Productivity was also confirmed as a decisive control regarding bankruptcy prediction. These results give strong support on those who focus and control closely efficiency and productivity when managing a firm.

5. CONCLUDING REMARKS

This dissertation focused on analyzing corporate bankruptcy, default and exit in Portugal. Completing this endeavor was a demanding and hard task nevertheless, immensely gratifying. The underlying purpose was trying to uncover the possible existence of variables, both idiosyncratic and(or) systemic, that could act as drivers and(or) predictors of corporate bankruptcy and(or) default. A particularly relevant feature would be the ability to study the period before and after the severe economic crisis that hit Portugal in 2008, disclosing in what measure this turmoil affected the corporate bankruptcy and default levels in the country. Being practically impossible to analyze the full set of companies in Portugal, choosing a sample that could, in the best way possible, represent the corporate sector in Portugal would be the starting objective. I was able to collect a sample of financial statements from a group of medium and large Portuguese companies due to the contributions and availability from the credit scoring service provider: **COFACE**. The collected data included: (1) full annual financial statements from 2006 to 2011; (2) information on headcount for the last year financial statements were available; (3) information on the presence of court actions – a dummy variable – particularly due to firms defaulting on schedule payments; (3) information on firms business exits – a dummy variable – that occurred inside the group of selected firms for the sample within the time period used for this dissertation; (4) information on firms bankruptcies – a dummy variable – that happened inside the group of selected firms for the sample within the time frame selected for this dissertation. From this data collection, I computed a group of financial and productivity indicators, for all observations in the sample. To this new treated sample, I added several Portuguese economic indicator for the years included in the study. Finally, I completed the dataset including information regarding firms' maturity and dimension.

I started the dissertation by reviewing the developments on the literature regarding credit scoring and default prediction. I presented a brief summary on the most updated theories and practices on these matters alongside with several papers on the empirical application on several types of companies and countries. I ended this review presenting a summary on firms Exits literature including a study regarding Portuguese case.

The dissertation proceeded detailing the selected data, namely describing the qualitative information included in the dataset, the financial ratios chosen for analysis, the accounting and computational procedures used for their extraction, supporting the several choices on the selected corporate ratios and systematic variables on similar studies and literature about bankruptcy. Following the presentation of the data in the dataset, I proceeded with the disclosure of a set of generic statistics with the purpose of detailing and identifying general characteristics and trends in the data. The selected sample included a group of 4.714 large medium-large Portuguese firms, excluding banking and financial sectors, that hold nearly 22% of employed people in Portugal. Highlight for the importance of construction sector both in dimension as in the part it represents in the bankruptcies group. It was mentioned that the number of bankruptcy events in the dataset is similar to other studies on bankruptcies supporting the relevance of the conclusions that would be produced in this dissertation. I presented summary statistics regarding dimension items of the firms in the dataset and for all selected financial ratios, these reveal the generic dimension and performance of the large Portuguese company. Finally I added generic statistics for the economic variables included in the thesis.

Entering in the driving purpose of the thesis, I proceeded with the econometric analysis, starting by analyzing and testing means on several indicators (ratios) for bankrupt and non-bankrupt firms. As logic anticipates there are consistent differences between the two groups of firms, particularly financial and productivity ratios. Bankrupt companies are clearly worse in profitability, more leveraged and less productive. All differences revealed to be statistically significantly different from zero. Going deeper in the analysis, I applied binary outcome models to the dataset with the objective of fitting a function of explanatory variables having as outcome the variable regarding bankruptcy. Starting with the brief description of the different available models and related options, together with the theoretical fundamentals and explanation for the reasons behind the choice of the model and options that would be used, the work went forward with the application of a Logit model with random-effects option.

Starting with a limited set of variables and growing in complexity, adding new variables in each step of the analysis, having as dependent variable the dummy for bankruptcy, the main conclusions can be described as: (1) profitability variables, surprisingly, revealed to be statistically non-significant for bankruptcy prediction. Due to the strongly positive correlation between profit ratios, increasing profitability information in the model did not add relevant information regarding bankruptcy prediction. This is somehow a surprising conclusion, contradicting results on other studies, see Bottazzi G., Grazi I., Secchi A. And Tamagni F. (2009) and Bernhardsen E. (2001). The nature of the corporate structure in Portugal allows that several firms remain in business even with negative profits, and this is influencing our dataset, which biases conclusions regarding profitability; (2) results on age are also contrary to other studies, see Castro, E. Christian, (2008). Two variables were tested, Age and Maturity (dummy), and, interpretation on results would lead us to consider that, advancing in age does not protect firms against bankruptcy probability. We saw that the means for these two variables are very similar between non-bankrupt and bankrupt firms and so, maturity did not prove a sufficient shield against bankruptcy in the crisis period that the Portuguese economy is facing since 2008; (3) ratios regarding leverage and liquidity disclose anticipated results, however only Solvency is statistically significant; (4) regarding systematic variables none is statistically significant with the exception of Investment, thought having a positive coefficient in the model, an economically absurd result implying that increasing investment in the economy would increase bankruptcy probability; (5) finally, **the most relevant variables for bankruptcy prediction are Assets Turnover and the dummy for the existence of court actions against a firm**. Assets Turnover has the most relevant coefficient from the group of corporate ratios, meaning that increasing the ratio would be a strong shield against bankruptcy. However, and by far, the most relevant variable in the model, is the dummy for court actions. These actions are indicating defaults from a firm, normally from scheduled payments on different stakeholders: Suppliers, Workers, Banks; naturally, if a firm starts defaulting, it can jeopardize severely its business operation, increasing bankruptcy risk.

The analysis went deeper in the dataset applying the same Logit model to particular issues. One first approach was to apply the model considering only firms belonging to the construction sector. Construction sector was probably the one that suffered the most with the economic turmoil that Portugal is crossing since 2008. Conclusions are mainly similar to the ones when all dataset was modeled and differences regard: (1) Liquidity, Solvency and Dimension are more relevant than in general model, particularly liquidity and Dimension that became statistically relevant but with different interpretations. Liquidity increased the negative coefficient, becoming a stronger shield against bankruptcy, meaning that if ratio is increased, bankruptcy probability lowers. In the Dimension control, the interpretation is the opposite which is, in some way, an odd conclusion disclosing that increasing book assets value would increase bankruptcy probability; this is certainly related with the kind of companies selected for the sample, we saw that dimension average was equivalent between bankrupt and non-bankrupt firms; (3) finally the dummy regarding the existence of court actions remains the most relevant variable in the model.

Next I applied the Logit model with the variables lagged by one year. Most of them lost statistical significance, only Dimension gained relevance, becoming statistically significant when data is lagged by one year.

The dataset includes information regarding business exit, allowing the possibility of analyzing exit decisions. The logic of the study was similar to the one applied in bankruptcy, first to compare and test means on some variables in the two groups and finally fitting a Logit model. Regarding mean differences particular attention to financial leverage, being much higher in companies that exited business. Solvency is other relevant result being lower in companies that exited business. Finally a Logit model was applied using as dependent variable the dummy for business exit. Comparing to generic model for bankruptcy prediction, most variables loose statistic significance, even court actions control, with the exception of Solvency and Age that are acting as shields against business exit probability.

The final approach regards testing default. The dummy variable regarding court charges was utilized as a proxy for default. Again the same logic as in prior cases: first to compare and test mean differences, secondly fitting a Logit model. All tested ratios reveled differences in means between the two groups, with the group having legal charges having the worst results. All tested means differences are statistically significant. After fitting the Logit model only three variables reveled statistic significance: Solvency, Dimension and Maturity. The intriguing result comes from dimension with a positive coefficient, but as detailed, this is certainly related with the bias in the dataset towards large companies and larger companies interact with more stakeholders and conflict probability is for sure higher.

This dissertation finalizes with hazard and survival-time analysis; the purpose is trying to understand the time to the occurrence of an event, in our case, bankruptcy, and, what are the variables that can influence the amount of time until the event; A Cox regression model was fitted using idiosyncratic variables. The results have some similarity with the ones from Logit models; profitability is now relevant for the model, particularly ROA and ROE which are statistically significant and their increase lowers hazard rate increasing survival-time for bankruptcy. Similar conclusions are obtained regarding Solvency, Financial Leverage, Dimension and Age. Assets Turnover and Court Actions control reveled the two most relevant variables in these models, the ones that exert stronger influence for hazard and survival-time for bankruptcy.

In sum, in this thesis, I thoroughly explore a detailed data set of medium and large Portuguese firms to examine the determinants of bankruptcy, default, and exit and finally proceed with a study regarding hazard and survival-time. Particularly relevant and a distinct feature from similar works, the study focused over a time frame during which the country presents two very distinct economic periods. I was able to identify variables that revealed significant for the purposes of the study namely, predicting hazard events and survival-time.

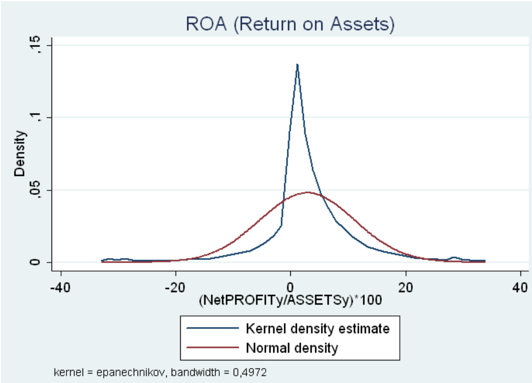
This dissertation was a demanding but extremely rewarding endeavor, from which I profited and will, for sure, profit in the future, both at personal and professional level.

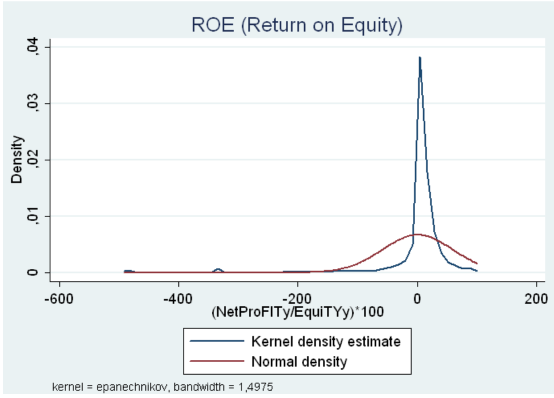
APPENDICES

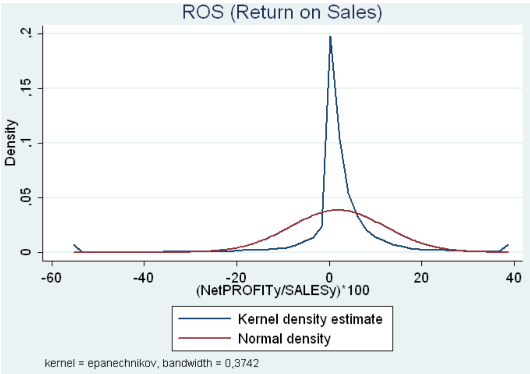
Appendix A: Variables description

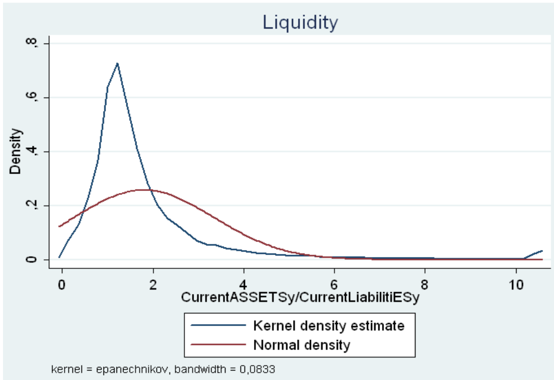
variable name	storage type	display format	value label		variable label
ID	int	%10.0g			ID
Year	int	%9.0g			Year for Observations
CAE_DESC	str100	%100s			Description for Portuguese Economical Sector
DATA_CONSTITU~O	str20	%20s			Company Start Date
StartYear	int	%10.0g			Company Start Year
IND_ACCOES_CI~S	long	%10.0g			Court action for defualting on Suppliers; = 1 if yes
IND_PROCESSOS~V	byte	%10.0g			Court Action (or) if Banckrupt; =1 if yes
IND_DISSOLUCOES	byte	%10.0g			Company Exit; = 1 if yes
ULT_ANOBAL	int	%10.0g			Last year Statutory Financial statements
ULT_ANOBAL_ANEM	long	%10.0g			ULT_ANOBAL_ANEM
Assets	double	%10.0g		*	In EUROS
NetProfit	double	%10.0g		*	In EUROS
ROA	double	%10.0g		*	(NetPROFITy/ASSETSy)*100
Equity	double	%10.0g			In EUROS
ROE	double	%10.0g		*	(NetProFITy/EquiTYy)*100
Sales	double	%10.0g			In EUROS
ROS	double	%10.0g		*	(NetPROFITy/SALESy)*100
CurrentAssets	double	%10.0g			In EUROS
CurrentLiabil~s	double	%10.0g			IN EUROS
Liquidity	double	%10.0g		*	CurrentASSETSy/CurrentLiabi > litiESy
EBIT	double	%10.0g			In EUROS
ErnPower	double	%10.0g		*	(EBITy/ASSETSy)*100
BusinessMargin	float	%9.0g		*	(EBITy/SALESy)*100
AssetsTurnover	float	%9.0g		*	SALESy/ASSETSy
FinancialDebt	double	%10.0g		*	In EUROS = Bank Loans + Bonds + ShareHolder Loans
InterestExpen~e	double	%10.0g		*	In EUROS = Interests on: Bank Loans and overdrafts + Bonds + ShareHolder Loans
InterestCover~e	float	%9.0g		*	(EBITy/InterestEXPENDITUREy >)
GDPVAR	double	%10.0g		*	GDPVAR
Investment	double	%10.0g		*	Percentage of GDP
Inflation	double	%10.0g		*	Inflation
Master_CAE	str1	%9s			Master_CAE
CAEdescription	str115	%115s			Description for Master CAE
YearsActivity	float	%9.0g		*	Firm's Age
Maturity	byte	%9.0g		*	1 if a firm is not mature (<= 10 years age)
Dimension	float	%9.0g			log(Assets)

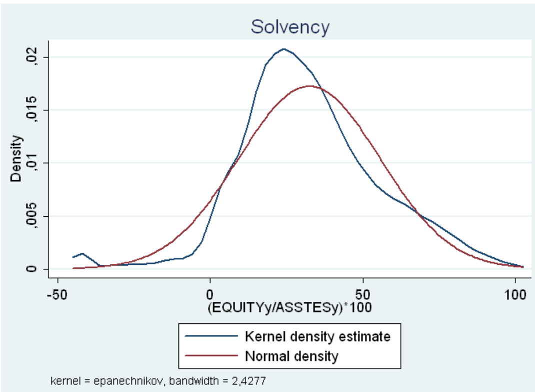
Appendix B: Kernel Density Plots

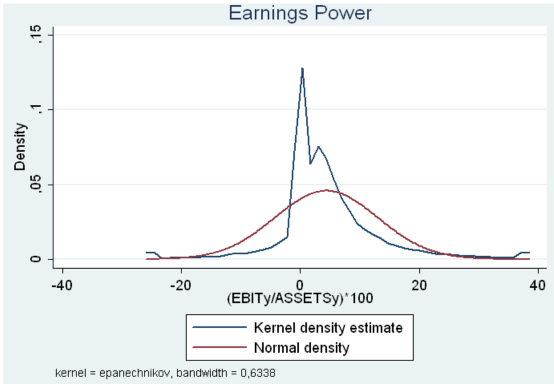


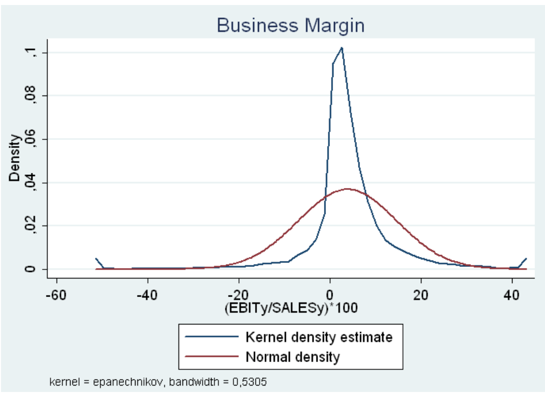


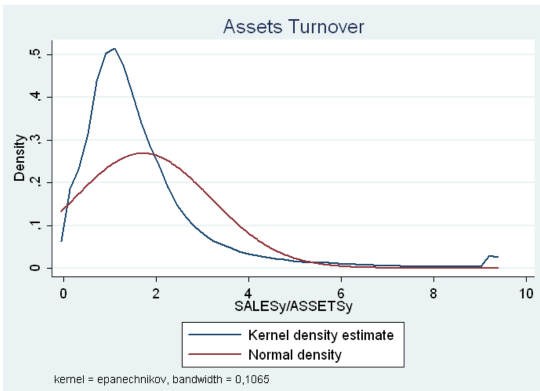


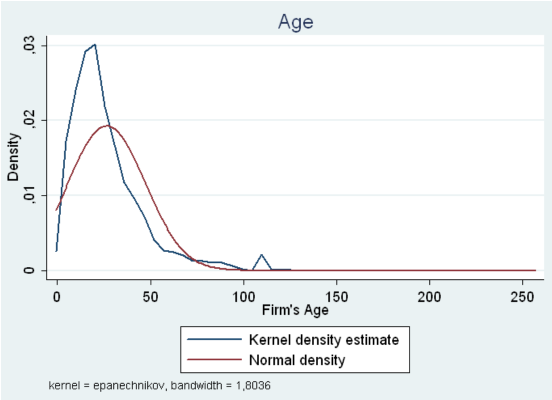


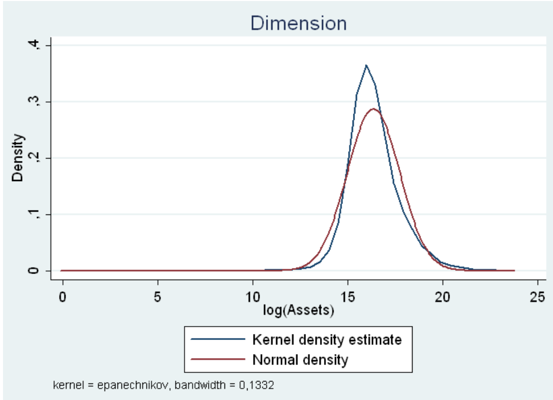









Age
Density
0
.01
.02
.03
0
50
100
150
200
250
Firm's Age
Kernel density estimate
Normal density
kernel = epanechnikov, bandwidth = 1,8036


Dimension
Density
0
.1
.2
.3
.4
0
5
10
15
20
25
log(Assets)
Kernel density estimate
Normal density
kernel = epanechnikov, bandwidth = 0,1332

Appendix C: Tests on Means Differences (ttest)

```
. ttest ROA, by(IND_PROCESSOS_INSOLV) unequal welch
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	25087	3,113618	,0522865	8,281591	3,011133	3,216102
1	1572	-1,009257	,1899029	7,529357	-1,381746	-,6367668
combined	26659	2,870504	,0508101	8,296055	2,770914	2,970095
diff		4,122874	,1969695		3,736564	4,509185

```
    diff = mean(0) - mean(1)                                      t =  20,9315
Ho: diff = 0                             Welch's degrees of freedom =  1817,88

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest ROE, by(IND_PROCESSOS_INSOLV) unequal welch
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	25085	1,937093	,3619291	57,32321	1,227691	2,646495
1	1572	-21,58075	2,158081	85,56459	-25,81378	-17,34773
combined	26657	,5502133	,3651526	59,61836	-,1655052	1,265932
diff		23,51784	2,18822		19,22588	27,80981

```
    diff = mean(0) - mean(1)                                      t =  10,7475
Ho: diff = 0                             Welch's degrees of freedom =  1660,65

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest ROS, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   24874    2,110851    ,0646153     10,1908    1,984201    2,237501
       1 |    1561   -2,221994    ,2891487    11,42412   -2,789155   -1,654833
---------+--------------------------------------------------------------------
combined |   26435    1,854994     ,063462    10,31818    1,730606    1,979383
---------+--------------------------------------------------------------------
    diff |            4,332845    ,2962805                3,751737    4,913953
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  14,6241
Ho: diff = 0                             Welch's degrees of freedom =  1719,63

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest Liquidity, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   24868    1,823583    ,0099259     1,56528    1,804128    1,843039
       1 |    1569    1,449848    ,0283085    1,121316    1,394322    1,505375
---------+--------------------------------------------------------------------
combined |   26437    1,801403    ,0095022    1,545012    1,782778    1,820028
---------+--------------------------------------------------------------------
    diff |            ,3737352    ,0299982                ,3149037    ,4325667
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  12,4586
Ho: diff = 0                             Welch's degrees of freedom =  1975,89

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest Solvency, by(IND_PROCESSOS_INSOLV) unequal welch
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	25085	33,16554	,1465116	23,20486	32,87837	33,45271
1	1572	20,71891	,4682827	18,56669	19,80039	21,63744
combined	26657	32,43155	,1417505	23,14356	32,15371	32,70938
diff		12,44663	,4906672		11,48432	13,40893

```
    diff = mean(0) - mean(1)                                      t =  25,3667
Ho: diff = 0                                 Welch's degrees of freedom =  1892,89

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest ErnPower, by(IND_PROCESSOS_INSOLV) unequal welch
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	25067	4,56626	,0553054	8,756261	4,457858	4,674662
1	1571	1,35833	,1811284	7,179176	1,003051	1,713609
combined	26638	4,377069	,0533294	8,703971	4,272541	4,481598
diff		3,207929	,1893837		2,836505	3,579354

```
    diff = mean(0) - mean(1)                                      t =  16,9388
Ho: diff = 0                                 Welch's degrees of freedom =  1875,76

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

. ttest BusinessMargin, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	24877	4,018944	,0680995	10,74095	3,885465	4,152423
1	1561	,8181443	,2947793	11,64658	,2399388	1,39635
combined	26438	3,829956	,0665608	10,82263	3,699493	3,960419
diff		3,200799	,3025432		2,607411	3,794188

diff = mean(0) - mean(1) t = 10,5796
Ho: diff = 0 Welch's degrees of freedom = 1730,87

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1,0000 Pr(|T| > |t|) = 0,0000 Pr(T > t) = 0,0000

. ttest AssetsTurnover, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	24877	1,740372	,0095047	1,499128	1,721742	1,759002
1	1561	1,239681	,0276942	1,094186	1,18536	1,294003
combined	26438	1,710809	,0091206	1,48299	1,692932	1,728686
diff		,5006905	,0292799		,4432673	,5581137

diff = mean(0) - mean(1) t = 17,1002
Ho: diff = 0 Welch's degrees of freedom = 1947,94

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1,0000 Pr(|T| > |t|) = 0,0000 Pr(T > t) = 0,0000

```
. ttest InterestCoverage, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	23058	1004,323	43,79732	6650,562	918,4769	1090,168
1	1514	188,1673	80,01817	3113,519	31,20901	345,1256
combined	24572	954,0353	41,41212	6491,539	872,865	1035,206
diff		816,1553	91,22014		637,2819	995,0286

```
    diff = mean(0) - mean(1)                                      t =   8,9471
Ho: diff = 0                          Welch's degrees of freedom =  2541,71

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest FinancialLeverage, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	21094	90,59064	3,657517	531,2095	83,42162	97,75965
1	1505	17,67464	4,257183	165,1546	9,323997	26,02529
combined	22599	85,73473	3,427815	515,3021	79,01598	92,45349
diff		72,916	5,612578		61,9125	83,91949

```
    diff = mean(0) - mean(1)                                      t =  12,9915
Ho: diff = 0                          Welch's degrees of freedom =  4377,39

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest YearsActivity, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   26460    26,45556    ,1284326    20,89153    26,20382    26,70729
       1 |    1824    26,92105    ,4110032    17,55325    26,11497    27,72714
---------+--------------------------------------------------------------------
combined |   28284    26,48557    ,1230392    20,69254    26,24441    26,72674
---------+--------------------------------------------------------------------
    diff |           -,4654971    ,4306025               -1,309928    ,3789339
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1,0810
Ho: diff = 0                             Welch's degrees of freedom =  2195,37

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,1399         Pr(|T| > |t|) = 0,2798          Pr(T > t) = 0,8601


. ttest Dimension, by(IND_PROCESSOS_INSOLV) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   25087    16,34655    ,0088356    1,399456    16,32924    16,36387
       1 |    1572    16,59677    ,0283574    1,124328    16,54115     16,6524
---------+--------------------------------------------------------------------
combined |   26659    16,36131    ,0084886    1,385986    16,34467    16,37795
---------+--------------------------------------------------------------------
    diff |           -,2502201     ,029702               -,3084723   -,1919678
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -8,4243
Ho: diff = 0                             Welch's degrees of freedom =  1890,12

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 1,0000
```

Appendix D: STATA® Correlation Matrix

```
. correlate IND_PROCESSOS_INSOLV IND_ACCOES_CIVEIS ROA ROE ROS Liquidity Solvency ErnPower BusinessMarg
> in AssetsTurnover InterestCoverage FinancialLeverage Maturity Dimension GDPVAR Investment Inflation
(obs=21407)
```

	IND_PR~V	IND_AC~S	ROA	ROE	ROS	Liquid~y	Solvency	ErnPower	Busine~n
IND_PROCES~V	1,0000								
IND_ACCOES~S	0,1709	1,0000							
ROA	-0,1220	-0,1201	1,0000						
ROE	-0,0973	-0,0746	0,6254	1,0000					
ROS	-0,0976	-0,0717	0,7135	0,4930	1,0000				
Liquidity	-0,0484	-0,0790	0,2397	0,1064	0,2078	1,0000			
Solvency	-0,1290	-0,1477	0,4873	0,2881	0,4040	0,4733	1,0000		
ErnPower	-0,0873	-0,1007	0,8064	0,4904	0,5328	0,1625	0,3476	1,0000	
BusinessMa~n	-0,0669	-0,0437	0,6414	0,4240	0,7659	0,1584	0,3092	0,6049	1,0000
AssetsTurn~r	-0,0867	-0,1009	0,0870	0,0648	-0,0085	-0,0753	-0,1083	0,1063	-0,0851
InterestCo~e	-0,0194	-0,0334	0,1607	0,0636	0,0998	0,0953	0,0966	0,1419	0,1072
FinancialL~e	-0,0359	-0,0191	0,0852	0,0362	0,0652	0,1262	0,1075	0,0572	0,0422
Maturity	-0,0385	-0,0365	-0,0618	-0,0470	-0,0889	-0,0999	-0,2102	-0,0426	-0,0711
Dimension	0,0456	0,2084	-0,0175	-0,0299	0,0583	0,0404	0,0719	-0,0547	0,1064
GDPVAR	0,0219	0,0078	0,0663	0,0666	0,0410	-0,0191	-0,0065	0,1970	0,0291
Investment	0,0408	0,0193	0,0592	0,0562	0,0258	0,0386	-0,0131	0,1051	0,0026
Inflation	-0,0196	-0,0076	-0,0009	0,0004	0,0000	-0,0451	0,0078	0,1991	0,0100

	Assets~r	Inter~ge	Financ~e	Maturity	Dimens~n	GDPVAR	Invest~t	Inflat~n
AssetsTurn~r	1,0000							
InterestCo~e	0,0137	1,0000						
FinancialL~e	0,0265	0,1258	1,0000					
Maturity	0,1066	-0,0011	0,0120	1,0000				
Dimension	-0,4968	0,0132	0,0178	-0,1092	1,0000			
GDPVAR	0,0150	0,0111	0,0052	-0,0420	-0,0461	1,0000		
Investment	0,0228	0,0040	0,0270	-0,0446	-0,0684	0,6057	1,0000	
Inflation	0,0246	-0,0069	-0,0057	-0,0250	-0,0201	0,3960	0,2155	1,0000

Appendix E: STATA® Panel Data Commands and Statistics

Panel identifier command *xtset* :

```
. xtset
       panel variable:  ID (strongly balanced)
        time variable:  Year, 2006 to 2011
                delta:  1 unit
```

Output indicates that data are available for all individuals in all periods (strongly balanced) and that the time variable increments by one.

```
. xtdescribe

      ID:  1, 2, ..., 4714                                      n =       4714
    Year:  2006, 2007, ..., 2011                                T =          6
           Delta(Year) = 1 unit
           Span(Year)  = 6 periods
           (ID*Year uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         6       6       6         6         6       6       6
```

Freq.	Percent	Cum.	Pattern
4714	100,00	100,00	111111
4714	100,00		XXXXXX

```
. xtsum ROA ROE ROS Liquidity Solvency FinancialLeverage Dimension
```

Variable		Mean	Std. Dev.	Min	Max	Observations
ROA	overall	2,870504	8,296055	-32,43607	33,33638	N = 26659
	between		6,673796	-31,96072	33,33638	n = 4714
	within		5,154661	-43,4228	53,39797	T-bar = 5,65528
ROE	overall	,5502133	59,61836	-489,796	99,38653	N = 26657
	between		39,64952	-316,0134	95,46993	n = 4714
	within		46,15037	-449,8933	341,6483	T-bar = 5,65486
ROS	overall	1,854994	10,31818	-54,83317	38,25615	N = 26435
	between		8,250696	-54,83317	38,25615	n = 4714
	within		6,573875	-75,71944	69,12481	T-bar = 5,60776
Liquid~y	overall	1,801403	1,545012	,0002316	10,49675	N = 26437
	between		1,2803	,0702923	10,49675	n = 4713
	within		,8890896	-5,354375	10,26209	T-bar = 5,60938
Solvency	overall	32,43155	23,14356	-42,59107	100	N = 26657
	between		21,27455	-42,59107	98,57919	n = 4714
	within		9,745807	-70,62517	139,3748	T-bar = 5,65486
Financ~e	overall	85,73473	515,3021	,0420512	4610,933	N = 22599
	between		486,4897	,2895885	4610,933	n = 4495
	within		351,0117	-3597,838	3926,72	T-bar = 5,02759
Dimens~n	overall	16,36131	1,385986	0	23,65477	N = 26659
	between		1,326434	8,087039	23,55379	n = 4714
	within		,4385665	4,021367	24,44835	T-bar = 5,65528

Appendix F: STATA® Panel Data xtlogit estimations

- **Xtlogit(1):**

```
Random-effects logistic regression              Number of obs      =     26657
Group variable: ID                              Number of groups   =      4714

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       5,7
                                                               max =         6

                                                Wald chi2(4)       =     74,96
Log likelihood  = -1483,5239                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,029813	,0122073	-2,44	0,015	-,053739	-,005887
Solvency	-,0288414	,0051923	-5,55	0,000	-,0390181	-,0186648
Maturity	-1,261795	,3234903	-3,90	0,000	-1,895824	-,6277651
Investment	,0895979	,0372579	2,40	0,016	,0165738	,162622
_cons	-10,95721	,8126072	-13,48	0,000	-12,54989	-9,364526
/lnsig2u	3,804523	,0412492			3,723676	3,88537
sigma_u	6,701032	,138206			6,435555	6,97746
rho	,9317365	,0026236			,9264114	,9367026

Likelihood-ratio test of rho=0: chibar2(01) = 8305,38 Prob >= chibar2 = 0,000

- **Xtlogit(2):**

```
Random-effects logistic regression              Number of obs      =     22599
Group variable: ID                              Number of groups   =      4495

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       5,0
                                                               max =         6

                                                Wald chi2(7)       =     80,78
Log likelihood  = -1412,7649                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0293058	,0137629	-2,13	0,033	-,0562805	-,002331
Solvency	-,0320195	,0060659	-5,28	0,000	-,0439084	-,0201305
Maturity	-1,185426	,3524037	-3,36	0,001	-1,876125	-,4947279
Investment	,1050658	,0474975	2,21	0,027	,0119725	,1981592
FinancialLeverage	-,0012507	,0005358	-2,33	0,020	-,0023008	-,0002006
Dimension	,2666771	,0769584	3,47	0,001	,1158413	,4175128
GDPVAR	-,0017036	,0531798	-0,03	0,974	-,1059341	,1025269
_cons	-15,64695	1,696715	-9,22	0,000	-18,97245	-12,32145
/lnsig2u	3,865635	,0434022			3,780568	3,950702
sigma_u	6,908948	,1499319			6,621249	7,209149
rho	,9355224	,002618			,9301971	,9404675

Likelihood-ratio test of rho=0: chibar2(01) = 7580,65 Prob >= chibar2 = 0,000

- **Xtlogit(3):**

```
Random-effects logistic regression              Number of obs      =     22366
Group variable: ID                              Number of groups   =      4485

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       5,0
                                                               max =         6

                                                Wald chi2(12)      =     81,33
Log likelihood  = -1409,3289                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0306479	,0240088	-1,28	0,202	-,0777042	,0164084
ROE	-,0003471	,0014551	-0,24	0,811	-,003199	,0025048
Solvency	-,0310797	,0063727	-4,88	0,000	-,04357	-,0185894
Maturity	-1,281075	,3826876	-3,35	0,001	-2,031129	-,5310214
Liquidity	-,0291382	,0938828	-0,31	0,756	-,2131451	,1548687
ErnPower	-,0022319	,0207381	-0,11	0,914	-,0428779	,038414
BusinessMargin	,0028805	,0104041	0,28	0,782	-,017511	,0232721
Investment	,1086017	,0477082	2,28	0,023	,0150955	,202108
FinancialLeverage	-,0012279	,0005174	-2,37	0,018	-,0022419	-,0002139
YearsActivity	-,0049697	,0061813	-0,80	0,421	-,0170849	,0071455
Dimension	,2730956	,0787165	3,47	0,001	,118814	,4273771
GDPVAR	-,0039203	,0535676	-0,07	0,942	-,1089109	,1010704
_cons	-15,65011	1,709844	-9,15	0,000	-19,00134	-12,29888
/lnsig2u	3,863599	,0435013			3,778338	3,94886
sigma_u	6,90192	,1501213			6,613871	7,202514
rho	,9353995	,0026287			,9300522	,9403643

Likelihood-ratio test of rho=0: chibar2(01) = 7509,17 Prob >= chibar2 = 0,000

- **Xtlogit(4):**

```
Random-effects logistic regression              Number of obs      =     21407
Group variable: ID                              Number of groups   =      4425

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4,8
                                                               max =         6

                                                Wald chi2(17)      =    486,46
Log likelihood  = -1161,0964                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0273954	,0324823	-0,84	0,399	-,0910595	,0362686
ROE	-,0005988	,0017503	-0,34	0,732	-,0040293	,0028317
ROS	-,0017483	,0198773	-0,09	0,930	-,0407072	,0372105
Liquidity	-,0544616	,1163557	-0,47	0,640	-,2825146	,1735913
Solvency	-,0279828	,0080619	-3,47	0,001	-,0437839	-,0121817
ErnPower	,0143843	,0264836	0,54	0,587	-,0375227	,0662913
BusinessMargin	,0019687	,018211	0,11	0,914	-,0337242	,0376616
AssetsTurnover	-,7570599	,1739993	-4,35	0,000	-1,098092	-,4160274
InterestCoverage	-9,24e-06	,0000316	-0,29	0,770	-,0000711	,0000527
FinancialLeverage	-,0011115	,0011021	-1,01	0,313	-,0032715	,0010485
GDPVAR	,0239584	,0701603	0,34	0,733	-,1135533	,16147
Investment	,0984545	,0581253	1,69	0,090	-,015469	,212378
Inflation	-,0890961	,0765778	-1,16	0,245	-,2391858	,0609936
YearsActivity	-,001529	,0081884	-0,19	0,852	-,017578	,0145201
Maturity	-1,227243	,4433564	-2,77	0,006	-2,096205	-,3582801
Dimension	-,1347315	,1136031	-1,19	0,236	-,3573895	,0879265
IND_ACCOES_CIVEIS	9,07094	,4344028	20,88	0,000	8,219526	9,922354
_cons	-18,54127	2,406214	-7,71	0,000	-23,25737	-13,82518
/lnsig2u	4,554672	,0461692			4,464181	4,645162
sigma_u	9,750668	,2250903			9,31933	10,20197
rho	,9665546	,0014925			,9635026	,9693595

Likelihood-ratio test of rho=0: chibar2(01) = 6827,82 Prob >= chibar2 = 0,000

- **XtLogit Construction:**

```
Random-effects logistic regression              Number of obs      =      1948
Group variable: ID                              Number of groups   =       386

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       5,0
                                                               max =         6

                                                Wald chi2(17)      =    185,05
Log likelihood  = -207,55545                    Prob > chi2        =    0,0000

------------------------------------------------------------------------------------
IND_PROCESSOS_INSOLV |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------------
   IND_ACCOES_CIVEIS |   14,33807   1,688056     8,49   0,000     11,02954     17,6466
                 ROA |  -,0885807   ,1470895    -0,60   0,547    -,3768709    ,1997094
                 ROE |  -,0208092   ,0113483    -1,83   0,067    -,0430514     ,001433
                 ROS |   ,1310495   ,0802071     1,63   0,102    -,0261536    ,2882525
           Liquidity |  -1,063433   ,3782905    -2,81   0,005    -1,804869   -,3219975
            Solvency |  -,0677808   ,0295813    -2,29   0,022    -,1257591   -,0098025
           ErnPower |   ,0991067   ,1212785     0,82   0,414    -,1385948    ,3368081
      BusinessMargin |   -,063804   ,0640001    -1,00   0,319    -,1892419    ,0616339
      AssetsTurnover |  -,8678032   ,9759383    -0,89   0,374    -2,780607    1,045001
    InterestCoverage |  -,0000513   ,0000402    -1,28   0,202    -,0001301    ,0000275
   FinancialLeverage |  -,0011389   ,0010102    -1,13   0,260     -,003119    ,0008411
              GDPVAR |   ,1688003   ,2802328     0,60   0,547    -,3804458    ,7180464
          Investment |   ,2242721   ,2373305     0,94   0,345    -,2408871    ,6894313
           Inflation |  -,3471554   ,2895223    -1,20   0,231    -,9146087    ,2202979
       YearsActivity |   ,0343485   ,0129305     2,66   0,008     ,0090052    ,0596919
            Maturity |   1,531219   1,440109     1,06   0,288    -1,291342     4,35378
           Dimension |   ,9663313   ,3179768     3,04   0,002     ,3431081    1,589554
               _cons |  -35,84991   8,126177    -4,41   0,000    -51,77693    -19,9229
---------------------+--------------------------------------------------------------
            /lnsig2u |   4,612143   ,1477781                      4,322503    4,901783
---------------------+--------------------------------------------------------------
             sigma_u |   10,03493   ,7414713                      8,681998    11,59868
                 rho |   ,9683635   ,0045273                      ,9581798    ,9761291
------------------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =  1527,21 Prob >= chibar2 = 0,000

.
```

- **XtLogit estimations with Crises dummy**

```
Random-effects logistic regression              Number of obs      =     21407
Group variable: ID                              Number of groups   =      4425

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4,8
                                                               max =         6

                                                Wald chi2(17)      =    504,16
Log likelihood  = -1122,7153                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
IND_ACCOES_CIVEIS	10,43762	,4931967	21,16	0,000	9,470973	11,40427
ROA	-,027656	,0343706	-0,80	0,421	-,0950211	,0397092
ROE	-,0006009	,0017906	-0,34	0,737	-,0041104	,0029087
ROS	,0012	,0194075	0,06	0,951	-,036838	,039238
Liquidity	-,0444819	,1214919	-0,37	0,714	-,2826016	,1936377
Solvency	-,0287106	,0083953	-3,42	0,001	-,0451652	-,0122561
ErnPower	,016595	,0283895	0,58	0,559	-,0390473	,0722374
BusinessMargin	,0016214	,017325	0,09	0,925	-,032335	,0355778
AssetsTurnover	-,8198756	,2492215	-3,29	0,001	-1,308341	-,3314104
InterestCoverage	-,0000119	,0000463	-0,26	0,798	-,0001026	,0000788
FinancialLeverage	-,0017904	,0008244	-2,17	0,030	-,0034062	-,0001745
GDPVAR	,0479831	,07926	0,61	0,545	-,1073636	,2033297
Inflation	-,081165	,0762853	-1,06	0,287	-,2306814	,0683515
YearsActivity	-,0017309	,0083261	-0,21	0,835	-,0180499	,014588
Maturity	-1,312688	,6687237	-1,96	0,050	-2,623362	-,0020133
Dimension	-,1988305	,1301591	-1,53	0,127	-,4539376	,0562765
Crisis	-,1661043	,2870675	-0,58	0,563	-,7287463	,3965376
_cons	-18,27498	2,308428	-7,92	0,000	-22,79941	-13,75054
/lnsig2u	4,847439	,0481546			4,753058	4,941821
sigma_u	11,28777	,2717791			10,76746	11,83321
rho	,9748296	,0011816			,972407	,9770445

Likelihood-ratio test of rho=0: chibar2(01) = 6930,97 Prob >= chibar2 = 0,000

Appendix G: STATA® xtlogit estimations Variables Lagged by one Year

```
Random-effects logistic regression              Number of obs      =     17966
Group variable: ID                              Number of groups   =      4366

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4,1
                                                               max =         5

                                                Wald chi2(17)      =    522,70
Log likelihood  =  -1015,726                    Prob > chi2        =    0,0000
```

IND_PROCESSOS_INSOLV	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA L1.	-,0223205	,0461305	-0,48	0,628	-,1127345	,0680936
ROE L1.	,0000874	,0032287	0,03	0,978	-,0062407	,0064155
ROS L1.	-,0032968	,0239946	-0,14	0,891	-,0503253	,0437317
Liquidity L1.	-,1458588	,1579509	-0,92	0,356	-,455437	,1637193
Solvency L1.	-,0392761	,012057	-3,26	0,001	-,0629074	-,0156449
ErnPower L1.	,0196637	,0381835	0,51	0,607	-,0551745	,0945019
BusinessMargin L1.	,0055289	,0229533	0,24	0,810	-,0394587	,0505165
AssetsTurnover L1.	-1,081952	,2448787	-4,42	0,000	-1,561905	-,6019981
InterestCoverage L1.	-,0000112	,0000543	-0,21	0,837	-,0001176	,0000952
FinancialLeverage L1.	-,0010605	,0008064	-1,32	0,188	-,0026409	,00052
GDPVAR L1.	,0254391	,155749	0,16	0,870	-,2798234	,3307017
Investment L1.	,1107979	,1491704	0,74	0,458	-,1815708	,4031665
Inflation L1.	-,1084676	,2679044	-0,40	0,686	-,6335505	,4166153
YearsActivity L1.	-,0140273	,0105504	-1,33	0,184	-,0347058	,0066512
Maturity L1.	-1,817883	,6410954	-2,84	0,005	-3,074407	-,5613593
Dimension L1.	-,3156724	,1530585	-2,06	0,039	-,6156616	-,0156833
IND_ACCOES_CIVEIS L1.	21,36102	1,00019	21,36	0,000	19,40068	23,32135
_cons	-33,15359	4,232766	-7,83	0,000	-41,44965	-24,85752
/lnsig2u	5,57776	,0511014			5,477603	5,677917
sigma_u	16,26279	,4155258			15,46843	17,09795
rho	,9877138	,0006201			,986437	,9888717

Likelihood-ratio test of rho=0: chibar2(01) = 6148,54 Prob >= chibar2 = 0,000

Appendix H: STATA® Summary Statistics and Tests on means on Exits

Summary statistics: mean
by categories of: IND_DISSOLUCOES (Company Exit; = 1 if yes)

IND_DISSOLUCOES	ROA	Solvency	Assets~r	Financ~e	YearsA~y	Dimens~n
0	2,870101	32,51462	1,710399	84,4681	26,64264	16,35922
1	2,888913	28,64298	1,729525	145,6319	20,57377	16,44032
Total	2,870504	32,43155	1,710809	85,73473	26,48557	16,36096

. ttest ROA, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	26087	2,870101	,0513773	8,298198	2,769398	2,970803
1	572	2,888913	,3430634	8,204882	2,215093	3,562733
combined	26659	2,870504	,0508101	8,296055	2,770914	2,970095
diff		-,0188119	,3468892		-,7000834	,6624596

diff = mean(0) - mean(1) t = -0,0542
Ho: diff = 0 Welch's degrees of freedom = 596,984

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0,4784 Pr(|T| > |t|) = 0,9568 Pr(T > t) = 0,5216

. ttest Solvency, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	26085	32,51462	,1435232	23,18022	32,23331	32,79594
1	572	28,64298	,8813254	21,07824	26,91194	30,37401
combined	26657	32,43155	,1417505	23,14356	32,15371	32,70938
diff		3,871644	,8929352		2,117996	5,625291

diff = mean(0) - mean(1) t = 4,3359
Ho: diff = 0 Welch's degrees of freedom = 601,785

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1,0000 Pr(|T| > |t|) = 0,0000 Pr(T > t) = 0,0000

. ttest AssetsTurnover, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	25871	1,710399	,0092288	1,484398	1,69231	1,728488
1	567	1,729525	,0595673	1,418402	1,612525	1,846524
combined	26438	1,710809	,0091206	1,48299	1,692932	1,728686
diff		-,0191254	,060278		-,1375094	,0992586

diff = mean(0) - mean(1) t = -0,3173
Ho: diff = 0 Welch's degrees of freedom = 593,588

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0,3756 Pr(|T| > |t|) = 0,7511 Pr(T > t) = 0,6244

. ttest FinancialLeverage, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	22131	84,4681	3,437762	511,4184	77,72984	91,20636
1	468	145,6319	31,05666	671,8584	84,60382	206,66
combined	22599	85,73473	3,427815	515,3021	79,01598	92,45349
diff		-61,16382	31,24635		-122,5608	,2331798

diff = mean(0) - mean(1) t = -1,9575
Ho: diff = 0 Welch's degrees of freedom = 478,562

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0,0254 Pr(|T| > |t|) = 0,0509 Pr(T > t) = 0,9746

```
. ttest YearsActivity, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   27552    26,64264    ,1250382    20,75483    26,39756    26,88772
       1 |     732    20,57377    ,6354009    17,19109    19,32634     21,8212
---------+--------------------------------------------------------------------
combined |   28284    26,48557    ,1230392    20,69254    26,24441    26,72674
---------+--------------------------------------------------------------------
    diff |            6,068869     ,647587                4,797671    7,340066
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   9,3715
Ho: diff = 0                             Welch's degrees of freedom =  788,838

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000


. ttest Dimension, by(IND_DISSOLUCOES) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   26087    16,35922    ,0086245     1,39298    16,34232    16,37613
       1 |     572    16,44032    ,0534161    1,277528    16,33541    16,54524
---------+--------------------------------------------------------------------
combined |   26659    16,36096    ,0085171    1,390631    16,34427    16,37766
---------+--------------------------------------------------------------------
    diff |           -,0810999    ,0541079               -,1873633    ,0251635
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1,4989
Ho: diff = 0                             Welch's degrees of freedom =  601,255

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,0672         Pr(|T| > |t|) = 0,1344          Pr(T > t) = 0,9328
```

Appendix I: STATA® Logit Model on Exits

```
Random-effects logistic regression              Number of obs      =     21407
Group variable: ID                              Number of groups   =      4425

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4,8
                                                               max =         6

                                                Wald chi2(16)      =     15,95
Log likelihood  = -536,07474                    Prob > chi2        =    0,4565
```

IND_DISSOLUCOES	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
IND_ACCOES_CIVEIS	-,6753947	,505292	-1,34	0,181	-1,665749	,3149594
ROA	,0385765	,0577145	0,67	0,504	-,0745418	,1516948
ROE	-,0006676	,0040479	-0,16	0,869	-,0086013	,007266
ROS	,0107779	,0375918	0,29	0,774	-,0629008	,0844565
Liquidity	,0384529	,1770862	0,22	0,828	-,3086298	,3855355
Solvency	-,028468	,0140378	-2,03	0,043	-,0559815	-,0009545
ErnPower	-,0185971	,0432426	-0,43	0,667	-,1033511	,0661569
BusinessMargin	-,0140295	,0312223	-0,45	0,653	-,0752241	,0471652
AssetsTurnover	-,2471506	,2130565	-1,16	0,246	-,6647336	,1704325
InterestCoverage	8,48e-06	,0000354	0,24	0,811	-,0000608	,0000778
FinancialLeverage	,0001497	,0003252	0,46	0,645	-,0004876	,000787
GDPVAR	,0108031	,1466044	0,07	0,941	-,2765363	,2981425
Investment	,1437542	,124309	1,16	0,248	-,0998869	,3873953
Inflation	-,1212642	,16879	-0,72	0,472	-,4520866	,2095582
YearsActivity	-,0374279	,0162322	-2,31	0,021	-,0692424	-,0056134
Dimension	-,0222252	,1981641	-0,11	0,911	-,4106197	,3661694
_cons	-29,77124	4,370161	-6,81	0,000	-38,3366	-21,20588
/lnsig2u	5,49657	,0416543			5,414929	5,578211
sigma_u	15,61583	,325233			14,99122	16,26646
rho	,9866885	,0005471			,9855724	,9877192

Likelihood-ratio test of rho=0: chibar2(01) = 3127,87 Prob >= chibar2 = 0,000

Appendix J: Default Analysis

- Tests On Mean Differences:

```
. ttest ROA, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   10413    4,215759     ,081669    8,333837    4,055672    4,375846
       1 |   16246    2,008253    ,0639898    8,156133    1,882826     2,13368
---------+--------------------------------------------------------------------
combined |   26659    2,870504    ,0508101    8,296055    2,770914    2,970095
---------+--------------------------------------------------------------------
    diff |            2,207506    ,1037522                2,004144    2,410868
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  21,2767
Ho: diff = 0                             Welch's degrees of freedom =  21845,6

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest ROE, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   10412    6,659013    ,5083552    51,87217    5,662539    7,655487
       1 |   16245   -3,365134    ,5004663    63,78737   -4,346103   -2,384165
---------+--------------------------------------------------------------------
combined |   26657    ,5502133    ,3651526    59,61836   -,1655052    1,265932
---------+--------------------------------------------------------------------
    diff |            10,02415    ,7133664                8,625908    11,42239
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  14,0519
Ho: diff = 0                             Welch's degrees of freedom =  25202,1

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest ROS, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   10308    2,921861    ,0959085    9,737431    2,733861     3,10986
       1 |   16127    1,173079    ,0836046    10,61712    1,009204    1,336953
---------+--------------------------------------------------------------------
combined |   26435    1,854994     ,063462    10,31818    1,730606    1,979383
---------+--------------------------------------------------------------------
    diff |            1,748782    ,1272327                1,499397    1,998166
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  13,7447
Ho: diff = 0                             Welch's degrees of freedom =  23319,3

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest Solvency, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	10412	37,22617	,233388	23,81473	36,76868	37,68365
1	16245	29,3585	,1739018	22,1648	29,01763	29,69937
combined	26657	32,43155	,1417505	23,14356	32,15371	32,70938
diff		7,867668	,2910529		7,297182	8,438154

```
    diff = mean(0) - mean(1)                                      t =  27,0317
Ho: diff = 0                                 Welch's degrees of freedom =  21028,4

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest BusinessMargin, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	10309	4,52762	,10265	10,42239	4,326406	4,728834
1	16129	3,384038	,0869932	11,04814	3,213522	3,554554
combined	26438	3,829956	,0665608	10,82263	3,699493	3,960419
diff		1,143582	,1345542		,8798468	1,407318

```
    diff = mean(0) - mean(1)                                      t =   8,4990
Ho: diff = 0                                 Welch's degrees of freedom =  22888,5

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest AssetsTurnover, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	10309	1,885405	,0161832	1,643132	1,853682	1,917127
1	16129	1,599215	,0107026	1,359225	1,578237	1,620193
combined	26438	1,710809	,0091206	1,48299	1,692932	1,728686
diff		,2861896	,0194021		,2481598	,3242194

```
    diff = mean(0) - mean(1)                                      t =  14,7505
Ho: diff = 0                                 Welch's degrees of freedom =    18978

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 0,0000
```

```
. ttest FinancialLeverage, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    8302     103,481    6,226033    567,2869    91,27641    115,6856
       1 |   14297    75,42982    4,033397    482,2735    67,52383     83,3358
---------+--------------------------------------------------------------------
combined |   22599    85,73473    3,427815    515,3021    79,01598    92,45349
---------+--------------------------------------------------------------------
    diff |            28,05117    7,418341                13,51033    42,59201
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3,7813
Ho: diff = 0                             Welch's degrees of freedom =  15179,8

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,9999         Pr(|T| > |t|) = 0,0002          Pr(T > t) = 0,0001

. ttest YearsActivity, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   11010    26,12371    ,2042615    21,43286    25,72332    26,52409
       1 |   17274    26,71622    ,1537221    20,20378    26,41491    27,01753
---------+--------------------------------------------------------------------
combined |   28284    26,48557    ,1230392    20,69254    26,24441    26,72674
---------+--------------------------------------------------------------------
    diff |           -,5925152    ,2556428               -1,093593   -,0914375
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2,3177
Ho: diff = 0                             Welch's degrees of freedom =  22427,7

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,0102         Pr(|T| > |t|) = 0,0205          Pr(T > t) = 0,9898

. ttest Dimension, by(IND_ACCOES_CIVEIS) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |   10413       15,99    ,0123047    1,255624    15,96588    16,01412
       1 |   16246    16,59874    ,0111458    1,420635    16,57689    16,62058
---------+--------------------------------------------------------------------
combined |   26659    16,36096    ,0085171    1,390631    16,34427    16,37766
---------+--------------------------------------------------------------------
    diff |           -,6087389    ,0166022               -,6412803   -,5761975
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t = -36,6661
Ho: diff = 0                             Welch's degrees of freedom =  24108,2

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0,0000         Pr(|T| > |t|) = 0,0000          Pr(T > t) = 1,0000
```

- **XtLogit for Default:**

```
Random-effects logistic regression              Number of obs      =     21407
Group variable: ID                              Number of groups   =      4425

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4,8
                                                               max =         6

                                                Wald chi2(16)      =    545,56
Log likelihood  = -2551,8533                    Prob > chi2        =    0,0000
```

IND_ACCOES_CIVEIS	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0534342	,0368606	-1,45	0,147	-,1256795	,0188112
ROE	-,0037263	,0029447	-1,27	0,206	-,0094979	,0020452
ROS	-,0191631	,0200886	-0,95	0,340	-,058536	,0202099
Liquidity	-,0977601	,1229218	-0,80	0,426	-,3386824	,1431623
Solvency	-,0727171	,0098961	-7,35	0,000	-,0921131	-,053321
ErnPower	-,0006075	,0329615	-0,02	0,985	-,0652109	,0639958
BusinessMargin	,0279256	,0186527	1,50	0,134	-,0086331	,0644843
AssetsTurnover	,2100608	,1247942	1,68	0,092	-,0345314	,4546529
InterestCoverage	-,0000246	,0000245	-1,00	0,315	-,0000727	,0000235
FinancialLeverage	,0002315	,0002313	1,00	0,317	-,0002218	,0006849
GDPVAR	,0008666	,154366	0,01	0,996	-,3016851	,3034184
Investment	,2648569	,1171525	2,26	0,024	,0352422	,4944716
Inflation	-,0580259	,1436939	-0,40	0,686	-,3396607	,223609
YearsActivity	-,0175626	,0124346	-1,41	0,158	-,0419339	,0068086
Maturity	-1,615108	,553008	-2,92	0,003	-2,698984	-,5312323
Dimension	2,046212	,1301363	15,72	0,000	1,791149	2,301274
_cons	-33,63375	3,732866	-9,01	0,000	-40,95003	-26,31746
/lnsig2u	4,484011	,0349516			4,415508	4,552515
sigma_u	9,412191	,1644856			9,095263	9,740161
rho	,9641935	,0012067			,9617518	,9664848

```
Likelihood-ratio test of rho=0: chibar2(01) =   2,1e+04 Prob >= chibar2 = 0,000
```

Appendix L: STATA® results on StSet

```
. stset Year, id(ID) failure(IND_PROCESSOS_INSOLV) origin(time StartYear) enter(time EnterYear) exit(ti
> me ULT_ANOBAL)

                id:  ID
     failure event:  IND_PROCESSOS_INSOLV != 0 & IND_PROCESSOS_INSOLV < .
obs. time interval:  (Year[_n-1], Year]
 enter on or after:  time EnterYear
 exit on or before:  time ULT_ANOBAL
    t for analysis:  (time-origin)
            origin:  time StartYear

------------------------------------------------------------------------------
    28284  total obs.
     5203  obs. end on or before enter()
      766  obs. begin on or after exit
------------------------------------------------------------------------------
    22315  obs. remaining, representing
     4711  subjects
     1283  failures in multiple failure-per-subject data
    22315  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =       255
```

```
. stdescribe

         failure _d:  IND_PROCESSOS_INSOLV
   analysis time _t:  (Year-origin)
             origin:  time StartYear
  enter on or after:  time EnterYear
  exit on or before:  time ULT_ANOBAL
                 id:  ID
```

Category	total	per subject mean	min	median	max
no. of subjects	4711				
no. of records	22315	4,736786	1	5	5
(first) entry time		21,76481	0	17	250
(final) exit time		26,50159	1	21	255
subjects with gap	0				
time on gap if gap	0	.	.	.	.
time at risk	22315	4,736786	1	5	5
failures	1283	,2723413	0	0	5

```
. stsum

         failure _d:  IND_PROCESSOS_INSOLV
   analysis time _t:  (Year-origin)
             origin:  time StartYear
  enter on or after:  time EnterYear
  exit on or before:  time ULT_ANOBAL
                 id:  ID

         |               incidence       no. of    |------ Survival time -----|
         | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
   total |        22315    ,057495          4711          8        15        27
```

Appendix M: STATA® StCox Results

- **StCox(1)**

Cox regression -- Breslow method for ties

No. of subjects =	4711	Number of obs =	21986
No. of failures =	1270		
Time at risk =	21986		
		LR chi2(3) =	744,13
Log likelihood =	-7086,9925	Prob > chi2 =	0,0000

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	,9758105	,0037593	-6,36	0,000	,9684701	,9832065
Solvency	,9782295	,0015333	-14,04	0,000	,975229	,9812393
AssetsTurnover	,6780594	,021418	-12,30	0,000	,637354	,7213645

Cox regression -- Breslow method for ties

No. of subjects =	4711	Number of obs =	21986
No. of failures =	1270		
Time at risk =	21986		
		LR chi2(3) =	744,13
Log likelihood =	-7086,9925	Prob > chi2 =	0,0000

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0244869	,0038525	-6,36	0,000	-,0320377	-,0169361
Solvency	-,022011	,0015674	-14,04	0,000	-,025083	-,018939
AssetsTurnover	-,3885204	,0315872	-12,30	0,000	-,4504301	-,3266107

- **StCox(2):**

Cox regression -- Breslow method for ties

No. of subjects =	4450	Number of obs =	18738
No. of failures =	1226		
Time at risk =	18738		
		LR chi2(7) =	680,81
Log likelihood =	-6679,6409	Prob > chi2 =	0,0000

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	,9853068	,0048873	-2,98	0,003	,9757743	,9949324
Solvency	,9775486	,001691	-13,13	0,000	,9742399	,9808686
AssetsTurnover	,6367748	,0259092	-11,09	0,000	,5879658	,6896356
ROE	,9993139	,0003856	-1,78	0,075	,9985585	1,00007
FinancialLeverage	,9995558	,0001711	-2,59	0,009	,9992205	,9998913
Dimension	,9399592	,025143	-2,31	0,021	,8919494	,9905532
YearsActivity	,9266658	,0196461	-3,59	0,000	,8889492	,9659826

Cox regression -- Breslow method for ties

No. of subjects =	4450	Number of obs =	18738
No. of failures =	1226		
Time at risk =	18738		
		LR chi2(7) =	680,81
Log likelihood =	-6679,6409	Prob > chi2 =	0,0000

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,0148022	,0049602	-2,98	0,003	-,0245239	-,0050804
Solvency	-,0227073	,0017299	-13,13	0,000	-,0260977	-,0193168
AssetsTurnover	-,4513392	,0406881	-11,09	0,000	-,5310865	-,371592
ROE	-,0006863	,0003859	-1,78	0,075	-,0014426	,0000699
FinancialLeverage	-,0004443	,0001712	-2,59	0,009	-,0007798	-,0001087
Dimension	-,0619188	,026749	-2,31	0,021	-,1143459	-,0094917
YearsActivity	-,0761623	,0212008	-3,59	0,000	-,1177152	-,0346094

- **StCox(3)**

```
Cox regression -- Breslow method for ties

No. of subjects =         4374                    Number of obs   =     17845
No. of failures =         1190
Time at risk    =        17845
                                                  LR chi2(13)     =   1095,90
Log likelihood  =   -6213,8808                    Prob > chi2     =    0,0000
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	,9789204	,0079933	-2,61	0,009	,9633785	,9947129
Solvency	,9810802	,0018458	-10,15	0,000	,9774692	,9847044
AssetsTurnover	,6256486	,0274177	-10,70	0,000	,5741539	,6817617
ROE	,9992313	,0003966	-1,94	0,053	,9984543	1,000009
FinancialLeverage	,9991855	,0003431	-2,37	0,018	,9985132	,9998583
Dimension	,8658311	,0246198	-5,07	0,000	,8188973	,9154549
YearsActivity	,9274144	,0200675	-3,48	0,000	,8889052	,967592
Liquidity	1,003521	,0280661	0,13	0,900	,9499926	1,060065
ROS	1,005069	,004826	1,05	0,292	,9956551	1,014573
ErnPower	1,004155	,0069377	0,60	0,548	,9906494	1,017846
BusinessMargin	1,000214	,0043809	0,05	0,961	,9916642	1,008837
InterestCoverage	,9999989	,0000116	-0,09	0,925	,9999762	1,000022
IND_ACCOES_CIVEIS	7,037094	,8656753	15,86	0,000	5,52945	8,955808

```
Cox regression -- Breslow method for ties

No. of subjects =         4374                    Number of obs   =     17845
No. of failures =         1190
Time at risk    =        17845
                                                  LR chi2(13)     =   1095,90
Log likelihood  =   -6213,8808                    Prob > chi2     =    0,0000
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ROA	-,021305	,0081654	-2,61	0,009	-,0373089	-,0053011
Solvency	-,0191011	,0018813	-10,15	0,000	-,0227885	-,0154137
AssetsTurnover	-,4689664	,0438229	-10,70	0,000	-,5548577	-,3830751
ROE	-,0007689	,0003969	-1,94	0,053	-,0015469	9,03e-06
FinancialLeverage	-,0008148	,0003434	-2,37	0,018	-,0014879	-,0001417
Dimension	-,1440654	,0284348	-5,07	0,000	-,1997966	-,0883341
YearsActivity	-,0753547	,0216381	-3,48	0,000	-,1177647	-,0329448
Liquidity	,0035144	,0279676	0,13	0,900	-,0513011	,0583299
ROS	,0050566	,0048016	1,05	0,292	-,0043544	,0144676
ErnPower	,0041469	,006909	0,60	0,548	-,0093946	,0176883
BusinessMargin	,0002139	,00438	0,05	0,961	-,0083707	,0087986
InterestCoverage	-1,09e-06	,0000116	-0,09	0,925	-,0000238	,0000216
IND_ACCOES_CIVEIS	1,951195	,123016	15,86	0,000	1,710088	2,192302

REFERENCES

Altman I. Eduard,1968, "*Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*", Journal of Finance, 23, 58-609.

Beaver, W., 1966, "*Financial Ratios as Predictors of Failure*", Journal of Accounting Research, 77-111.

Bernhardsen E., 2001, "A Model of Bankruptcy Prediction", Norges Bank, Working Paper.

Bharath T. S. and Shumway T., 2004, "*Forecasting Default with the KMV-Merton Model*", University of Michigan, December 2004.

Bhattacharjee, A., Higson, C., Holly, S., Kattuman, P., 2009, "Macro Economic Instability and Business Exit: Determinants of Failures and Acquisitions of Large UK Firms", Economica, 76, 108-131.

Black F. and Scholes M., 1973, "*The Pricing of Options and Corporate Liabilities*", Journal of Political Economy, 81, 637-654.

Bonfim, D., 2009, "*Credit risk drivers: evaluating the contribution of firm level information and of macroeconomic dynamics*", Working Paper, Economic Research Department, Bank of Portugal.

Borio, C.**,** Furfin, C. and Lowe, P., 2001, "Procyclicality of the financial system and financial stability: issues and policy options", BIS papers No.1, March.

Bottazzi G., Grazi I., Secchi A. And Tamagni F., 2009, "Financial and Economic Determinants of Firm Default", Working Paper, LEM, Scuola Superiore Sant'Anna, Università di Pisa, Italia.

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

Campbel l, J., J. Hilscher, and J. Szilagyi, 2008, "In Search of Distress Risk", The Journal of Finance, 63, 2899-2938.

Cameron A. Colin and Triverdi Pravin K., 2009, "*Microeconometrics Using Stata – Rivised Edition*", STATA ® Press.

Castro, E. Christian, 2008, "*Estimating a financial distress rating system for Spanish firms with a simple hazard model*", Draft, June 2008.

CFA® Program Curriculum, 2011, "*Financial Reporting and Analysis*", Vol. 3, Reading 35, 307-319.

Chamberlain, Gary., 1980, "Analysis of Covariance With Qualitative Data." Review of Economic Studies, Vol. XLVII(1), No. 146, pp. 225-238.

Chava S. and J. A. Jarrow, 2004, "Bankruptcy Prediction with Industry Effects", Review of Finance, 8, 537-569.

Cleves M., Gould W., Gutierrez R., Marchenko Y., 2008, “An Introduction to Survival Analysis Using STATA”, STATA® Press.

Crosbie P. and Kocagil A., 2003, “*Modeling Default Risk Methodology*”, © Moody’s KMV Company.

Cox, D. R. 1972. “*Regression models and life-tables*.” Journal of the Royal Statistical Society, Series B 34: 187–220.

Duffie D., Saita L. and Wang K., 2007, “Multi-period corporate default prediction with stochastic covariates”, Journal of Financial Economics, 83, 635-665.

Eklund, T., Larsen, K. and Bernhardsen, E., 2001, “Model for analyzing credit risk in the enterprise sector”, Norges Bank Economic Bulletin Q3 01.

Gujarati, Damodar., 2003, " *Basic Econometrics* (4th ed.)” . McGraw−Hill.

Hillegeist, S. A., D. P. Cram, E. K. Keating, and K. G. Lundstedt, 2004, “*Assessing the Probability of Bankruptcy*”, Review of Accounting Studies, 9, 5-34.

López-Garcia P. and Puente S., 2006, “Business demography in Spain: determinants of firm survival”, Banco de España, Working Paper, No 0608.

M. Miller and F. Modigliani, 1958, “*The Cost of Capital, Corporation Finance, and Theory of Investment*”, The American Economic Review, 48, 261.297.

Mata, J., Antunes, A., Portugal, P., 2011, “Borrowing Patterns, Bankruptcy and Voluntary Liquidation”, Bank of Portugal.

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

Morris, R., 1997, “Early *Warning Indicators of Corporate Failure*”, Ashgate Publishing Ltd. Harris.

Ohlson, J., 1980, “*Financial ratios and the probabilistic prediction of bankruptcy*”, Journal of Accounting Research, Spring, 109-131.

Rosch, D., 2004, “*Correlations and business cycles of credit risk: evidence from bankruptcies in Germany*”, Financial Markets and Portfolio Management, 17, No.3, 309-331.

Saunders A. and Allen L., 2010, *Credit Risk – Measurement In and Out of the Financial Crisis, Third Edition*, John Wiley & Sons, Inc (Wiley Finance).

Shumway, T., 2001, “*Forecasting Bankruptcy More Accurately: A Simple Hazard Mod*el”,The Journal of Business, 74, 101-124.

Thornhill S. and Amit R., 2003, “Learning about Failure: Bankruptcy, Firm Age, and the Resource-Based View”, Organization Science, 14, No 5, 497-509.